PROBABILITIES ON NON-WELL-FOUNDED SETS

ANDREW SCHUMANN

HOLNY MEYERA, 2007

Aczel is the founder of non–well–founded set theory

Aczel, P.: Non–Well–Founded Sets. CSLI Lecture Notes No. 14, Stanford, 1988.

Now there are many axiomatizations of set teory without the foundation axiom.

A binary relation R on a set S is well-founded if there is no infinite sequence b0 , b1 , b2,... of elements of S such that R(bn+1, bn) for each n = 0,1, ... If there is such a sequence then R is said to be non-well-founded, and such a sequence is called a descending sequence for R.

R is said to be circular if there is a finite sequence b0,...,bk such that b0 = bk and R(bn+1, bn) for each n = 1, ..., k. Such a sequence is called a cycle in R.

If a relation R is circular then it is non-well-founded In other words, if R is well-founded then it is non-circular.

We might characterize this cyclical nature of time by means of unfolding ``streams" of weeks and seasons; they unfold without end but with a cyclic pattern to their nature.

week = (Su,(M,(Tu,(W,(Th,(Fr, (Sat, week)))))))

seasons = (spring, (summer, (fall, (winter, seasons))))

Cycles occur not only in the physical world around us, but also in the biological and psychological world within.

The instance of circularity in reasoning:

This proposition is false

The interest in non-well-founded phenomena is mainly motivated by some developments in computer sciences. Indeed, in this area, many objects and phenomena do have non-well-founded features: self-applicative programs, self-reference, graph circularity, looping processes, transition systems, paradoxes in natural languages, etc.

Some others like strings, streams, and formal series are potentially infinite, and can only be approximated by partial and progressive knowledge. Also, it is natural to use universes containing adequate non-well-founded sets as frameworks to give semantics for these objects or phenomena. Moreover, it is often not relevant to use the classical principles of definition and reasoning by induction to define and reason about these objects.

Denying the foundation axiom in number systems implies setting the non-Archimedean ordering structure. Remind that Archimedes' axiom affirms the existence of an integer multiple of the smaller of two numbers which exceeds the greater: for any positive real or rational number y, there exists a positive integer n such that $y = 1/n$ or $ny = 1$. The informal sense of Archimedes' axiom is that anything can be measured by a ruler.

The negation of Archimedes' axiom says that there exist infinitely small numbers (or infinitesimals), i.e., numbers that are smaller than all real or rational numbers of the open interval (0, 1) and as well as infinitely large integers that are greater than all positive integers.
Robinson applied this idea into modern mathematics and developed so-called non-standard analysis.

Within the framework of non-standard analysis there were obtained many interesting results. There exists also a different version of mathematical analysis in that Archimedes' axiom is rejected, namely, p-adic analysis. In this analysis, one investigates the properties of the completion of the field **Q** of rational numbers with respect to the p-adic (non-Archimedean) metric. This completion is called the field **Q**p of p-adic numbers. In **Q**p there are infinitely large integers.

In the case of non-Archimedean number systems we survey non-well-founded sets. For example, some their sets are obtained by the circular membership relation (that is, some sets of infinitesimals satisfy this property). It follows from this that the negation of Archimedes' axiom allows non-well-founded phenomena to be regarded too and provides new theoretical framework of consideration in the meantime.

In the standard way, probabilities are defined on an algebra of subsets. Recall that an algebra **A** of subsets $A \subseteq X$ consists of the following: (1) union, intersection, and difference of two subsets of X; (2) $\varnothing$ and X. Then a finitely additive probability measure is a nonnegative set function P( ) defined for sets A$\in$ **A** that satisfies the following properties:

1. $P(A) \geq 0$ for all A$\in$ **A**,
2. P (X) = 1 and P($\varnothing$) = 0,
3. if A$\in$ **A** and B$\in$ **A** are disjoint, then P(A B) = P(A) + P(B). In particular P($\neg$ A) = 1 – P(A) for all A$\in$ **A**.

It is possible also to set probabilities on an algebra F(X) of fuzzy subsets $A \subseteq X$ that consists of the following: (1) union, intersection, and difference of two fuzzy subsets of X; (2) $\varnothing$ and X. In this case a finitely additive probability measure is a nonnegative set function P( ) defined for sets $A \in F(X)$ that runs the non-Archimedean set V and satisfies the following properties:

1. P (A) ≥ 0 for all $A \in F(X)$,
2. P (X) = 1 and P($\varnothing$) = 0,
3. if $A \in F(X)$ and $B \in F(X)$ are disjoint, then P(A B) = P(A) + P(B).
4. P(¬ A) = 1 – P(A) for all $A \in F(X)$, where 1 is the largest member of V and 0 is the least member of V.

This probability measure is called fuzzy probability.

The main originality of fuzzy probabilities is that conditions 3, 4 are independent. As a result, in a probability space some Bayes' formulas do not hold in the general case.

A probability space <X, F(X), P> will say to be non-Archimedean. As we see it is a particular case of fuzzy probability space and non-Archimedean probability measure is a particular case of fuzzy probabilities.

Let S be a number system (e.g., the field of rational numbers). Remind that non-Archimedean extension of S is $^*S = S^\omega/U$ where U is Frechet filter. This means that $^*S$ consists of infinite tuples of the form $\langle x_0, x_1, \ldots \rangle$, where $x_i \in S$ for any $i = 0,1,..$

A non-Archimedean Bayesian network N consists of the following

1. V is a set included variables $v_1^i, \ldots, v_N^i$ of various order $i \in \omega$ and variables $v_1^\omega, \ldots, v_N^\omega$ of ω-order.
2. A is a union of (1) a set of i-order arc towers ($i \in \omega$), which together with V constitutes an i-order dag G_i over variables $v_1^1, \ldots, v_N^1$ at the first level, over variables $v_1^i, \ldots, v_N^i$ at the i-th level, etc., and (2) a set of ω-order arc towers, which together with V constitutes an ω-

order dag G_ω over variables $v_1^\omega, \ldots, v_N^\omega$.

3. P is a set of i-order conditional probabilities $P_i(v_j^i \mid \pi_{v_j^i})$ of the all i-order variables v_j^i given their respective i-order parents $\pi_{v_j^i}$ ($i \in \omega$) and of ω-order conditional probabilities $P_\omega(v_j^\omega \mid \pi_{v_j^\omega})$ of the all ω-order variables v_j^ω given their respective ω-order parents $\pi_{v_j^\omega}$.

Also we have a multihierarchical (more precisely, infinitely hierarchical) Bayesian network. For instance, we can consider i-order variables as i-tuples of first-order variables and ω-order variables as infinite tuples of first-order variables.

The main idea of non-Archimedean Bayesian networks is that we can define multihierarchical structures and consider joint distributions for different levels $i=1,2,\ldots$ Principles of setting an infinite hierarchy of Bayesian networks depend on practical aims.