

A Surprise-Based Qualitative Probability Calculus II

Zina Ibrahim, Ahmed Y. Tawfik and Alioune Ngom

{ibrahim,atawfik,angom}@uwindsor.ca

School of Computer Science, University of Windsor
401 Sunset Avenue, Windsor, Ontario N9B 3P4 Canada

Abstract

This paper is a continuation of the study of surprise as a base for constructing qualitative calculi for representing and reasoning about uncertain knowledge. Here, we further elaborate on κ^{++} , a qualitative ranking function which we developed in (Ibrahim, Tawfik, and Ngom 2009b) and which constructs qualitative ranks for events by obtaining the order of magnitude abstraction of the degree of surprise associated with them. Having formulated surprise propagation rules via order of magnitude conditional operators, we commit this paper to show some of the properties that κ^{++} possesses which make it an improvement on κ , the ranking function capturing the order of magnitude of probability.

Introduction and Overview of the κ^{++} Function

Several proposals exist for using statistical measures other than the p -value to model surprise (Bayarri and Berger 1997). Those measures are motivated by the observation that the inverse of a probability of an event is not a good measure of how 'unlikely' it is as it does not offer any indication of how comparable the number is with the existing model (Bayarri and Berger 1997). Taking this observation to qualitative ranking functions for modeling surprise, we developed κ^{++} (Ibrahim, Tawfik, and Ngom 2009b); a function aimed at improving ranking provided by the known κ model, which captures the order of magnitude abstraction of the probabilistic inverse of an event and uses the resulting number to rank events as surprising or non-surprising.

κ^{++} captures the order of magnitude abstraction of the Weaver surprise index (Weaver 1948), which numerically defines surprise as the ratio of the expected value of the probability to the probability of the event that actually occur and is denoted by $\mathcal{W}(w)$ for an event w . $\mathcal{W}(w)$ produces numbers from zero to infinity. Values between zero and one corresponding to a likely outcome (no surprise) while values greater than one indicate a surprise.

$\kappa^{++}(w) \in \mathbb{Z} \cup \{+\infty\}$ for all values w of a variable W is defined to be of the same order of magnitude of the Weaver index of the value w of W as given in Equation 1 below

Copyright © 2010, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

(Ibrahim, Tawfik, and Ngom 2009b) where ϵ is a small number greater than zero and less than one.

$$\epsilon < \frac{\mathcal{W}(w)}{\epsilon^{\kappa^{++}(w)}} \leq 1 \quad (1)$$

As a result of the abstraction, positive κ^{++} values indicate a surprise, negative values belong to events that are anticipated while a zero κ^{++} value indicates the occurrence of the event is neither surprising nor anticipated.

Moreover, κ^{++} obeys the rules of conjunction, disjunction and conditionality by replacing multiplication and addition by their order of magnitude counterparts of addition and minimum respectively.

Initial Evaluation of the κ^{++} Function

Having given a brief overview of κ^{++} , this section is concerned with showing its effectiveness by comparing it to κ with respect to the three parameters we define below.

1. Semantics In contrast to the κ function where having a κ rank of zero is universal to all non-surprising events, $\kappa^{++}(w) = 0$ is the threshold that separates surprising events from expected ones. This gives non-surprising events a better representativeness because essentially, an event with a negative κ^{++} rank is anticipated (as opposed to the surprising positive κ^{++} ranks), with the event being considered more expected as the smaller the negative κ^{++} value is.

Hence, κ^{++} stands out as offering a richer semantics that enables one to speak of degrees of anticipation in addition to degrees of surprise, providing better assignment of ranks to complements of variables in addition to variables themselves as Example 1 demonstrates.

Example 1. Consider two variables $W_1, W_2 \in \Omega$. Let $\kappa^{++}(w_1) = +2$ and $\kappa^{++}(w_2) = +5$. It is possible to induce that $\neg w_2$ is more believable than $\neg w_1$ if the zero rank is considered the threshold between surprising and anticipated events (hereby adhering by the rule $\kappa^{++}(W) + \kappa^{++}(\neg W) = 0$), resulting an assignment of $\kappa^{++}(\neg w_1) = -2 =$ and $\kappa^{++}(\neg w_2) = -5$. Notice that the κ calculus would have assigned a value of zero for both $\kappa(\neg w_1)$ and $\kappa(\neg w_2)$, providing no useful way for their comparison.

In addition to a better representativeness of the various belief states of the variables is that with the κ^{++} calcu-

lus, one is now able to $\kappa^{++}(\neg w_1)$ and $\kappa^{++}(\neg w_2)$ in addition to $\kappa^{++}(w_1)$ and $\kappa^{++}(w_2)$. This is not permissible in the κ calculus as deductive closure mandates that $\kappa(W) \vee \kappa(\neg W) = 0$, assigning both $\neg w_1$ and $\neg w_2$ κ ranks of zero. As a result, one is unable to predict which is the more anticipated (or believed) of the two nor propagate the epistemic state of the complements.

2. Domain Independence One important feature of the κ^{++} calculus is that the number it assigns is computed while taking into account the expected value of probability of the distribution (as it is part of computing the Weaver index). As a result, the rank assigned to the event is not designated with respect to the absolute 1 but instead with respect to the average probability expected for the domain. This entails that the κ^{++} rank assigned to the event takes into account the relative distribution of surprise and non-surprise in the domain, which makes its value of use other than in the ordinal sense if looked at from outside the domain. This is obviously not the case for κ whose probabilistic interpretation does not take into account the way the probabilities are distributed among the events under consideration, causing its values to be of a mere ordinal nature.

Example 2. Given two universes Ω_1, Ω_2 each containing a set of events, let $W_1 \in \Omega_1$ and $W_2 \in \Omega_2$ be two events. Let $\kappa(w_1) = 2, \kappa^{++}(w_1) = 1, \kappa(w_2) = 2$ and $\kappa^{++}(w_2) = 3$.

Examining the κ values associated with the outcomes of the events W_1 and W_2 does not provide an indication of their relative surprise relative to each other because $\kappa(w_1)$ is only meaningful within Ω_1 and $\kappa(w_2)$ is only meaningful in Ω_2 . Hence, despite the fact that these two outcomes have equal κ values, no conclusions can be made about whether or not they are equally surprising in their respective universes.

In contrast, one is able to deduce that within their respective distributions, w_2 is more surprising than w_1 by examining their κ^{++} values, where one can clearly see that w_2 is more surprising than w_1 not because of the value, but because one is certain that the values are computed with respect to the average (expected) surprise of Ω_1 and Ω_2 respectively, rendering w_2 more surprising than w_1 .

3. Better Ranking

It is well accepted that κ assumes an infinitesimal ϵ for its abstraction (Goldszmidt and Pearl 1996). When a non-infinitesimal ϵ is used, the rankings given by κ tend to deviate from those given by probability. Moreover, an infinitesimal κ gives trivial rankings and is therefore not usable.

We investigated how good the ranking provided by κ^{++} compared that provided by κ using different values of ϵ . The results are shown in Figure 1. The Comparison is made for extreme and non-extreme values of probability ranging between 0.05 and 0.45. The range is chosen so that only positive κ^{++} values are generated in order to compare with the non-signed κ values. The curves are Bezier interpolation of discrete values performed to demonstrate the difference between the values of κ and κ^{++} in terms of 1) the slope of the curve, which indicate the speed of degeneration to the trivial (zero) value 2) the range assumed by the function. Decreasing the value of ϵ causes the rank provided by κ to

quickly degenerate to the trivial case (zero), while the slope of the convergence to zero is less in the case of κ^{++} . Moreover, the larger the value of ϵ , the more trivial the values of κ compared to those of κ^{++} . As a result, κ^{++} provides a more useful and less trivial ranking than κ .

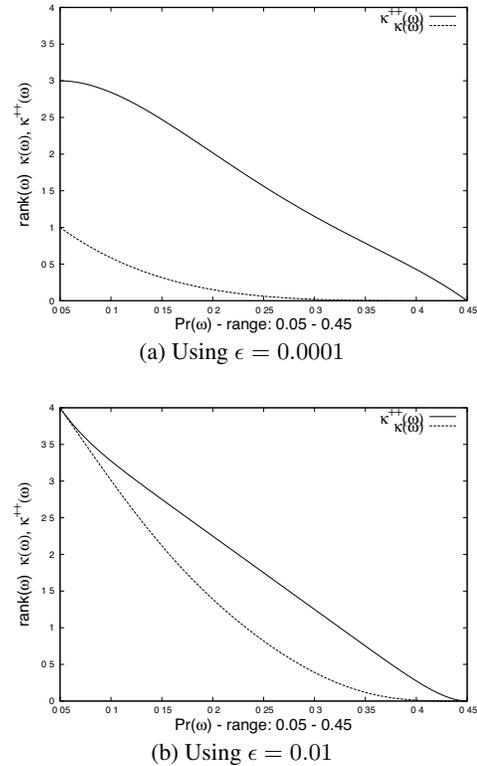


Figure 1: κ^{++} 's sensitivity to ϵ values compared to κ

Current Work

Part of our current work investigates incorporating κ^{++} as a strength indicator *Qualitative Probabilistic Networks* and comparing its power to κ 's in resolving conflicts in the highly coarse abstraction of Bayesian Networks (Ibrahim, Tawfik, and Ngom 2009a).

References

- Bayarri, M., and Berger, J. O. 1997. Measures of surprise in bayesian analysis. In *Duke University*.
- Goldszmidt, M., and Pearl, J. 1996. Qualitative Probabilities for Default Reasoning, Belief Revision, and Causal Modeling. *Artificial Intelligence* 84:145–151.
- Ibrahim, Z.; Tawfik, A.; and Ngom, A. 2009a. Surprise-based qualitative probabilistic networks. In *ECSQARU*, 228–239.
- Ibrahim, Z.; Tawfik, A.; and Ngom, A. 2009b. A surprise-based qualitative probability calculus. In *FLAIRS*, 573–574.
- Weaver, W. 1948. Probability, Rarity, Interest and Surprise. *Scientific Monthly* 67:390–392.