Semirings Modeling Confidence and Uncertainty in Speech Recognition

Günther J. Wirsching

January 3, 2011

Introduction

When looking up the word 'confidence' in the *Oxford Dictionaries Online*, one gets two meanings:

1. The feeling or belief that one can have faith in or rely on someone or something.
2. The telling of private matters or secrets with mutual trust.

As usual in speech recognition, in this paper I restrict attention to part of the first meaning, namely to "the belief that one can rely on something".

To get a feeling for this concept, assume that we have an automatic speech recognition system which is configured with a language model giving Bayesian prior probabilities to certain word sequences. If a user says something to the system, the system will produce a set of possible phrases the user may have said. Ideally, each phrase comes equipped with a score that can be interpreted as Bayesian posterior probability of that phrase, given the language model and the utterance. This Bayesian posterior probability measures the "degree of belief that a user uttered a specific phrase", whence it is a measure for *confidence*.

This paper studies the mathematical structures of this confidence measure and connects it to a reasonable measure for uncertainty. A thorough analysis of logical operations leads to semiring structures for these measures. The two arising semirings turn out to be isomorphic, and also isomorphic to the well-known max-plus algebra [1]. In addition, it is proved that, for any two of these semirings, the set of semiring isomorphisms between them is a one-parameter family.

1 Logical analysis of confidence and uncertainty

In automatic speech recognition, one often is confronted with the problem that the automaton is not quite sure about what it understood. Assume that a recognition result r consists of a set of phrases $\phi_1, \ldots, \phi_n$ such that each phrase comes with its Bayesian posterior probability

$$p_n = p_{\phi_n}(r) \in [0, 1]$$

such that $\sum p_i \leqslant 1$. If $\sum p_i < 1$, adjoin an additional empty phrase ϕ_0 meaning "nothing understood" with probability

$$p_0 = 1 - \sum_{i=1}^{n} p_i.$$

Here an example which occurs every day in Bavaria's bakeries. The client says something, and the recognition result r of the salesperson may consist of the following three phrases:

ϕ_1: "Zwei Semmeln, bitte." ("Two buns, please."), ($p_1 = 0.4$),

ϕ_2: "Drei Semmeln, bitte." ("Three buns, please."), ($p_2 = 0.5$),

ϕ_0: "..." ($p_0 = 0.1$).

In this setting, the *confidence* of a phrase ϕ in a recognition result r can be defined by

$$c_\phi(r) := \left\{ \begin{array}{l} \text{The degree of belief that phrase } \phi \text{ had been said,} \\ \text{given recognition result } r. \end{array} \right\} = p_\phi(r).$$

There are two natural logical operations on the set of possible recognition results: the exclusive disjunction, denoted by "*x-or*", and the conjunction, which we denote by "&". We derive some formulas which describe how these logical operations affect the confidence measure.

First consider *x-or*, which refers to a decision: Assume that you have two recognition results r_1 and r_2, and that you are to decide which one results from a recognition of an uttered phrase ϕ. Then, clearly, you would take the one in which ϕ has the higher confidence. Hence, it is reasonable to set

$$\begin{aligned} c_\phi(r_1 \text{ x-or } r_2) &:= \left\{ \begin{array}{l} \text{The degree of belief that phrase } \phi \text{ had been said,} \\ \text{given two mutually exclusive recognition results } r_1 \text{ and } r_2. \end{array} \right\} \\ &= \max\{c_\phi(r_1), c_\phi(r_2)\}. \end{aligned} \tag{1}$$

Next consider the conjunction "&", fix a phrase ϕ, and suppose that we have two recognition results r_1 and r_2. We are looking for the degree of belief that at least one of the recognition results r_1 and r_2 comes from recognizing phrase ϕ. To get this confidence of "$r_1 \,\&\, r_2$", denote for the moment by A_1 the event that r_1 comes from recognizing ϕ, and by A_2 the event that r_2 comes from recognizing ϕ. If we assume that events A_1 and A_2 are stochastically independent, the probability of $A_1 \cup A_2$ can be computed as follows:

$$\begin{aligned} p(A_1 \cup A_2) &= p(A_1) + p(A_2) - p(A_1)p(A_2) \\ &= p_\phi(r_1) + p_\phi(r_2) - p_\phi(r_1)p_\phi(r_2). \end{aligned}$$

Hence, the total confidence that some phrase ϕ had been intended, when two stochastically independent recognition results are given, is

$$\begin{aligned} c_\phi(r_1 \,\&\, r_2) &:= \left\{ \begin{array}{l} \text{The degree of belief that phrase } \phi \text{ had been said,} \\ \text{given two stochastically independent recognition results } r_1 \text{ and } r_2. \end{array} \right\} \\ &= c_\phi(r_1) + c_\phi(r_2) - c_\phi(r_1) \cdot c_\phi(r_2). \end{aligned} \tag{2}$$

Let us now shift attention to *uncertainty*, which we define as the probabilistic complement of confidence, which is, for a phrase ϕ and a recognition result r, the quantity

$$\begin{aligned} u_\phi(r) &:= 1 - c_\phi(r) = 1 - p_\phi(r) \\ &= \left\{ \begin{array}{l} \text{The uncertainty about phrase } \phi \text{ being said,} \\ \text{given recognition result } r. \end{array} \right\}. \end{aligned}$$

Here the *x-or* reads

$$\begin{aligned} u_\phi(r_1 \text{ x-or } r_2) &:= \left\{ \begin{array}{l} \text{The uncertainty about phrase } \phi \text{ being said,} \\ \text{given two mutually exclusive recognition results } r_1 \text{ and } r_2. \end{array} \right\} \\ &= \min\{u_\phi(r_1), u_\phi(r_2)\}, \end{aligned} \tag{3}$$

and the "&" is

$$\begin{aligned} u_\phi(r_1 \,\&\, r_2) &:= \left\{ \begin{array}{l} \text{The uncertainty about phrase } \phi \text{ being said,} \\ \text{given two stochastically independent recognition results } r_1 \text{ and } r_2. \end{array} \right\} \\ &= u_\phi(r_1) \cdot u_\phi(r_2). \end{aligned} \tag{4}$$

2 The algebraic structures

Recall that a *monoid* is an algebraic structure $\mathbb{M} = (M, \circ, e)$ consisting of a set M, a binary operation

$$\circ : M \times M \to M,$$

and a special element $e \in M$, subject to two conditions:

(M.1) $\circ$ is *associative*, i.e., $\forall\, x, y, z \in M :\ (x \circ y) \circ z = x \circ (y \circ z)$.

(M.2) e is a *neutral element* for $\circ$, i.e., $\forall\, x \in M :\ e \circ x = x \circ e = x$.

A monoid $\mathbb{M}$ is called *commutative* iff $\forall\, x, y \in M :\ x \circ y = y \circ x$. Note that it is an immediate consequence from the two-sided definition of neutrality (M2) that a binary operation admits at most one neutral element.

A *semiring* is an algebraic structure

$$\mathbb{S} = (S, \oplus, 0, \otimes, 1)$$

consisting of a set S, two binary operations $\oplus$ and $\otimes$ on S, and two special elements $0, 1 \in S$, subject to the following four conditions:

(S.1) $(S, \oplus, 0)$ is a commutative monoid.

(S.2) $(S \setminus \{0\}, \otimes, 1)$ is a monoid.

(S.3) The two laws of distributivity hold:

$$\begin{aligned} &\forall\, x, y, z \in S :\quad x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z), \\ &\forall\, x, y, z \in S :\quad (x \oplus y) \otimes z = (x \otimes z) \oplus (y \otimes z). \end{aligned}$$

(S.4) $\forall\, x \in S :\ 0 \otimes x = 0$. "0 annihilates S."

A semiring $\mathbb{S}$ is called *idempotent* if, in addition to (S.1)–(S.4), its **first** operation is idempotent:

(S.5) $\forall x \in S :\ x \oplus x = x$.

A semiring $\mathbb{S}$ is called *commutative* if, in addition to (S.1)–(S.4), its **second** operation is commutative:

(S.6) $\forall\, x, y \in S :\ x \otimes y = y \otimes x$.

Clearly, in a commutative semiring, it suffices to postulate only one of the distributivity laws, as the other then follows by commutativity. In contrast to the definition of the algebraic structure *ring*, for a semiring it is necessary to postulate the annihilation axiom (S.4), because the monoid structure with respect to the first binary operation $\oplus$ does not suffice to prove annihilation.

Note that axiom (S.2) implies that $0 \neq 1$. This will turn out to be an important point to be taken into account when constructing a semirng.

Given two (not necessarily commutative) semirings

$$\mathbb{A} = (A, \oplus_{\mathbb{A}}, 0_{\mathbb{A}}, \otimes_{\mathbb{A}}, 1_{\mathbb{A}}) \qquad \text{and} \qquad \mathbb{B} = (B, \oplus_{\mathbb{B}}, 0_{\mathbb{B}}, \otimes_{\mathbb{B}}, 1_{\mathbb{B}}),$$

a *homomorphism* $\phi : \mathbb{A} \to \mathbb{B}$ is a map $\phi : A \to B$ satisfying

(H.1) $\phi(0_{\mathbb{A}}) = 0_{\mathbb{B}}$.

(H.2) $\phi(1_{\mathbb{A}}) = 1_{\mathbb{B}}$.

(H.3) $\forall\, x, y \in A :\ \phi(x \oplus_{\mathbb{A}} y) = \phi(x) \oplus_{\mathbb{B}} \phi(y)$.

(H.4) $\forall\, x, y \in A :\ \phi(x \otimes_{\mathbb{A}} y) = \phi(x) \otimes_{\mathbb{B}} \phi(y)$.

If ϕ happens to be bijective, it follows that its inverse ϕ^{-1} is also a homomorphism of semirings. A bijective homomorphism of semirings is called an *isomorphism* of semirings.

3 Three isomorphic commutative, idempotent semirings

In this section, it is shown that the two binary operations *exclusive disjunction* and *conjunction* on confidences yield a semiring structure. Indeed, there are three distinct semirings, the *confidence semiring*, the *uncertainty semiring*, and the *max-plus-algebra*, and it is proved here that these are pairwise isomorphic. Moreover, it is shown that the set of isomorphisms between any two of them is a one-parameter family. The families of isomorphismus are constructed explicitly.

The construction of a semiring modeling uncertainty starts with the unit interval $[0,1]$, where a number $u \in [0,1]$ should express the *uncertainty* about the validity of some assertion. According to (3), *x-or* is modeled by the binary operation min, and, according to (4), the multiplication of numbers in $[0,1]$ models logical *conjunction*.

In this algebraic setting, the point $1 \in [0,1]$ would be a neutral element both with respect to min, and with respect to multiplication. Hence, in order to come to a semiring, we have to adjoin a further element ∞ and appropriate conventions to make it the 0 of the semiring. It turns out that the conventions

$$\forall u \in [0,1]: \quad \min\{\infty, u\} = \min\{u, \infty\} = u, \quad u \cdot \infty = \infty \cdot u = \infty,$$

ensure both neutrality with respect to min and the annihilation axiom. Commutativity, associativity, and the distrubitivity law

$$\forall\, u, v, w \in [0,1] \cup \{\infty\}: \quad \min\{u \cdot v, u \cdot w\} = u \cdot \min\{v, w\}$$

are consequences of elementary algebra of real numbers. Now we are ready to define the *uncertainty semiring* as the algebraic structure

$$\mathbb{U} := ([0,1] \cup \{\infty\}, \min, \infty, \cdot, 1).$$

To construct a semiring which models confidence, start again with the unit interval $[0,1]$, and assume that a number from this interval expresses the degree of belief that, given a feature and an utterance, the utterance intended to convey a value to the given feature. As described in (1) and (2), exclusive or and conjunction lead to the binary operations

$$\begin{aligned} c \sqcup d &:= \max\{c, d\} && \text{for exclusive or,} \\ c \sqcap d &:= c + d - cd && \text{for conjunction.} \end{aligned}$$

Similar to the structure modeling uncertainty, the number $0 \in [0,1]$ is neutral with respect to both operations, whence, in order to get a semiring, we have to include an additional zero, namely $-\infty$. By definition, the *confidence semiring* is the algebraic structure

$$\mathbb{T} := (\{-\infty\} \cup [0,1], \sqcup, -\infty, \sqcap, 0); \tag{5}$$

here "$\mathbb{T}$" stands for "trust". To see that this is indeed a semiring, consider the involution map

$$\iota : [0,1] \cup \{\infty\} \to \{-\infty\} \cup [0,1], \qquad \iota(c) := 1 - c. \tag{6}$$

This map is clearly bijective, and it fulfills $\iota(\infty) = -\infty$ and $\iota(0) = 1$, and

$$\forall\, c, d \in [0,1] \cup \{\infty\}: \quad \begin{cases} \iota(\min\{c,d\}) = \max\{\iota(c), \iota(d)\} = \iota(c) \sqcup \iota(d), \\ \iota(c \cdot d) = 1 - cd = (1-c) + (1-d) - (1-c)(1-d) \\ \qquad\qquad = \iota(c) + \iota(d) - \iota(c)\iota(d) = \iota(c) \sqcap \iota(d), \end{cases}$$

where, in case $c = \infty$ or $d = \infty$, the arithmetic rule $(-\infty) + \infty = \infty + (-\infty) = -\infty$ is to be applied. This proves both that the algebraic structure (5) is indeed a semiring, and that $\iota : \mathbb{U} \to \mathbb{T}$ is an isomorphism of semirings.

Finally, the *max-plus algebra* is the idempotent, commutative semiring

$$\mathbb{M} = (\{-\infty\} \cup [0,\infty], \max, -\infty, +, 0);$$

note that in this semiring, the above mentioned arithmetic rule is just the annihilation axiom.

Theorem 1 *The uncertainty semiring $\mathbb{U}$, the confidence semiring $\mathbb{T}$, and the max-plus algebra $\mathbb{M}$, are pairwise isomorphic. Moreover, the sets of isomorphisms are as follows:*

$$\mathrm{Iso}(\mathbb{U},\mathbb{M}) = \left\{x = -\alpha \ln u \mid \alpha > 0\right\}, \qquad \mathrm{Iso}(\mathbb{M},\mathbb{U}) = \left\{u = e^{-\alpha x} \mid \alpha > 0\right\}, \tag{7}$$

$$\mathrm{Iso}(\mathbb{T},\mathbb{M}) = \left\{x = -\beta \ln(1-c) \mid \beta > 0\right\}, \qquad \mathrm{Iso}(\mathbb{M},\mathbb{T}) = \left\{c = 1 - e^{-\beta x} \mid \beta > 0\right\}, \tag{8}$$

$$\mathrm{Iso}(\mathbb{U},\mathbb{T}) = \left\{c = 1 - u^{\gamma} \mid \gamma > 0\right\}, \qquad \mathrm{Iso}(\mathbb{T},\mathbb{U}) = \left\{u = (1-c)^{\gamma} \mid \gamma > 0\right\}, \tag{9}$$

where it is understood that $u \in \mathbb{U}$, $c \in \mathbb{T}$, and $x \in \mathbb{M}$.

Proof Suppose that $\phi : \mathbb{U} \to \mathbb{M}$ is a semiring homomorphism. By definition, ϕ satisfies (H.1)–(H.4); in particular,

$$\forall\, u, v \in]0,1] : \quad \phi(\min\{u,v\}) = \max\{\phi(u), \phi(v)\}, \quad \text{and} \quad \phi(u \cdot v) = \phi(u) + \phi(v).$$

The first equation implies that $\phi : [0,1] \to [0,\infty]$ is non-increasing, while the second equation is just the functional equation of the logarithm. Denoting be $e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2.7182$ the Eulerian number, and fixing the number

$$\alpha := \phi\left(\frac{1}{e}\right) \in [0,\infty[,$$

we conclude from the functional equation that $\phi(u) = -\alpha \ln(u)$ whenever $u = \left(\frac{1}{e}\right)^{\rho}$ for some rational number ρ. As the set of rational powers of $\frac{1}{e}$ is a dense set in $[0,1]$, monotonicity of ϕ implies that

$$\forall u \in]0,1] : \quad \phi(u) = -\alpha \ln u.$$

As, by (H.1) and (H.2), ϕ must satisfy

$$\phi(0) = \infty, \qquad \phi(\infty) = -\infty,$$

it follows that ϕ is a bijective map $[0,1] \cup \infty \to \{-\infty\} \cup [0,\infty]$ iff $\alpha > 0$, whence (7).

To see (8), observe that a semiring homomorphism $\psi : \mathbb{T} \to \mathbb{M}$ is an isomorphism iff $\psi \circ \iota : \mathbb{U} \to \mathbb{M}$ is an isomorphism, where ι is the involution map (6). Consequently, $\psi : \mathbb{T} \to \mathbb{M}$ is an isomorphism iff there is a real number $\beta > 0$ such that

$$\forall c \in [0,1[: \quad \psi(c) = -\beta \ln(1-c).$$

To determine the isomorphisms $\varrho : \mathbb{T} \to \mathbb{U}$, fix some real number $\alpha > 0$, and put

$$\phi_\alpha : \mathbb{U} \to \mathbb{M}, \quad \phi_\alpha(u) := -\alpha \ln u.$$

Then $\psi := \phi_\alpha \circ \varrho$ is an isomorphism $\mathbb{T} \to \mathbb{M}$. This implies by (8) that there is a positive real number β such that

$$\forall c \in [0,1[: \quad \psi(c) = -\beta \ln(1-c).$$

Using $\varrho = \phi_\alpha^{-1} \circ \psi$, we get

$$\forall c \in [0,1[: \quad \varrho(c) = e^{\alpha\beta \ln(1-c)} = (1-c)^{\alpha\beta}.$$

Putting $\gamma := \alpha\beta$, this gives the second equation of (9). The first equation is obtained from this by solving $u = (1-c)^{\gamma}$ for c, giving $c = 1 - u^{(1/\gamma)}$. As γ runs through the positive reals, also $1/\gamma$ runs through the positive reals, leading to the first equation of (9). ∎

References

[1] Peter Butkovič. *Max-linear System: Theory and Algorithms*. Springer, 2010.