UNIVERSITÄT AUGSBURG

The confidence-probability semiring

Günther J. Wirsching
Markus Huber
Christian Kölbl

Report 2010-04 April 2010

institut für informatik

INSTITUT FÜR INFORMATIK
D-86135 AUGSBURG

1 Introduction

The *confidence-probability semiring* serves as a modeling of uncertain information. Possible applications are the representation of dialogue systems and weighted automata [2] with the explicit use of confidence-numbers.

The *confidence* will be modeled using the *max-plus-algebra*, which is also called *arctic semiring*:

$$\begin{aligned}\mathbb{M}\mathrm{ax} &= (\mathbb{R}_+ \cup \{\pm\infty\}, \oplus_{\mathbb{M}\mathrm{ax}}, 0_{\mathbb{M}\mathrm{ax}}, \otimes_{\mathbb{M}\mathrm{ax}}, 1_{\mathbb{M}\mathrm{ax}}) \\ &= (\{-\infty\} \cup [0,\infty], \max, -\infty, +, 0)\,,\end{aligned}$$

The *probabilities* are modeled by the *min-plus-algebra*, which is also known by *tropical semiring*:

$$\begin{aligned}\mathbb{M}\mathrm{in} &= (\mathbb{R}_+ \cup \{\infty\}, \oplus_{\mathbb{M}\mathrm{in}}, 0_{\mathbb{M}\mathrm{in}}, \otimes_{\mathbb{M}\mathrm{in}}, 1_{\mathbb{M}\mathrm{in}}) \\ &= ([0,\infty], \min, \infty, +, 0).\end{aligned}$$

It will be shown below that each of these algebras is isomorphic to a semiring whose elements can be interpreted as probabilities.

The confidence-probability semiring will be constructed as some kind of „semidirect product“ of the arctic and the tropical semiring. The analogy to the semidirect product in group theory is that the operations on the product are not just component-wise operations on the underlying semirings, as it would be on the direct product, but an additional operation of the one semiring on the other semiring by means of semiring endomorphisms is used. In this case we define an operation of the max-plus-algebra on the min-plus-algebra.

The two operations $\oplus$ and $\otimes$ will be construced in a way that allows for information theoretical interpretation. The $\oplus$-addition combines information from incompatible sources, e. g., different results from an N-best speech recognizer decoding one utterance. The $\oplus$-operation consequently leads to a decision between two confidence-probability pairs: the one with the higher confidence is kept, and in case of equal confidences the one with the higher probability is kept.

In contrast, the $\otimes$-operation is to be interpreted as fusion of two confidence-probability-pairs: the confidences are summed up, and the probabilities of the product can be interpreted – towards the limit of high confidences – as a Bayesian Posterior if one perceives the initial probabilities as

Bayesian Prior and Likelihood. On the other hand, if one of the confidence-probability pairs has low confidence, the other one is barely changed by fusion.

This construction of the operations $\oplus$ and $\otimes$ have one downer: In order to have semiring structure on the product, it is necessary to exclude ∞ from the max-plus-algebra, because otherwise the distributive law could be broken.

The probabilistic semiring

To admit a probabilistic interpretation, we consider the *probabilistic semiring*

$$\mathbb{P} = \left([0,1] \cup \{-\infty\}, \max, -\infty, \uplus, 0\right),$$

whose multiplication is the *probabilistic sum*

$$p \uplus q := p + q - pq.$$

The following mapping will be important below:

$$\mu : \mathbb{M}\mathrm{ax} \to \mathbb{P}, \quad \mu(x) := 1 - e^{-x}, \quad \mu(-\infty) = -\infty.$$

It is a semiring-isomorphism, because it is obviously bijective, and it fulfills

$$\begin{aligned} \mu(\max\{x,y\}) &= 1 - e^{-\max\{x,y\}} \\ &= 1 - \min\left\{e^{-x}, e^{-y}\right\} \\ &= \max\{\mu(x), \mu(y)\}, \\ \mu(-\infty) &= -\infty, \end{aligned}$$

and

$$\begin{aligned} \mu(x+y) &= 1 - e^{-(x+y)} \\ &= 1 - e^{-x} \cdot e^{-y} \\ &= \left(1 - e^{-x}\right) + \left(1 - e^{-y}\right) - \left(1 - e^{-x}\right)\left(1 - e^{-y}\right) \\ &= \left(1 - e^{-x}\right) \uplus \left(1 - e^{-y}\right) \\ &= \mu(x) \uplus \mu(y) \\ \mu(0) &= 0. \end{aligned}$$

So, the probabilistic semiring is isomorphic to the max-plus-algebra.

The Bayesian-Semiring and the Min-Plus-Algebra

With the structure

$$\mathbb{H} = ([0,1], \max, 0, \cdot, 1)\,,$$

another semiring is given which we call *Bayesian-Semiring* because its multiplication appears to be well-suited for modeling the transition from a Bayesian Prior to the Posterior. The letter $\mathbb{H}$ is adopted to avoid confusion with the *Boolean Semiring*.

In speech recognition devices, in general negative logarithms of probabilities, so called „log-probs", are used in stead of probabilities. These result from „normal" probabilities through the mapping:

$$\ell : [0,1] \to [0,\infty], \quad \ell(p) := -\log p, \quad \ell(0) := \infty.$$

It follows from the fact that ℓ is order preserving, and from the functional equation of the logarithm, that ℓ is a semiring-isomorphism from the Bayesian-Semiring to the min-plus-algebra,

$$\ell : \mathbb{H} \xrightarrow{\cong} \mathbb{M}\mathrm{in}\,.$$

The operation of the probabilistic semiring on the Bayesian semiring

In his diploma thesis [1], Huber considers pairs (c, π), consisting of a *confidence* $c \in [0,1]$ and a *probability vector*

$$\pi = (p_1, \ldots, p_n) \in \Delta^n_+$$

taken from the *open standard simplex*

$$\Delta^n_+ := \left\{ (p_1, \ldots, p_n) \in \mathbb{R}^n \;\middle|\; \text{all } p_i > 0, \text{and } \sum_{i=1}^n p_i = 1 \right\}$$

to model uncertain information. Therein it is postulated that there will not occur zero-probabilities because it has technical advantages on the one hand and no apparent practical disadvantages on the other hand — particularly in practical applications in speech recognition as considered by Huber, it is usefull to allow for a certain amount of uncertainity at all times.

Huber constructs, based on the ideas of Römer and Wirsching, an „update-function“ as binary operation $\circ$ over the set

$$\mathbb{S}^n := [0,1] \times \Delta^n_+.$$

To define $\circ$, first consider the *standardization-function*

$$\mathcal{N} : (\mathbb{R}_+)^n \setminus \{0\} \to \Delta^n := \left\{ (p_1, \ldots, p_n) \in (\mathbb{R}_+)^n \;\middle|\; \sum_{i=1}^n p_i = 1 \right\},$$
$$\mathcal{N}(p_1, \ldots, p_n) := \frac{1}{p_1 + \ldots + p_n}(p_1, \ldots, p_n),$$

and define

$$(c, \pi) \circ (s, \lambda) := \left(c \uplus s, \mathcal{N}\left((\pi^c \cdot \lambda^s)^{1/(c \uplus s)} \right) \right). \tag{1}$$

Here the *probabilistic sum* $c \uplus s = c + s - cs$ shows up again, and exponation and multiplication of a probability vector are defined component-wise:

$$\pi^c \cdot \lambda^s := (p_1^c \lambda_1^s, \ldots, p_n^c \lambda_n^s)\ .$$

The update-function $\circ$ should model the fusion of two pieces of information to the same topic, each consisting of a confidence and a probability vector. It fulfills the following set of postulations:

1. $\circ$ is continuous as a function $\left(]0,1] \times \Delta^n_+\right) \times \left(]0,1] \times \Delta^n_+\right) \to \mathbb{S}^n$
2. $\circ$ is commutative: the order should not play a role by the fusion of pieces of information.
3. $\circ$ is associative: it should be unimportant in which way the pieces of information are combined.
4. $\circ$ is isotone in the first component: the confidence does not fall if new information is added — but the probability-vector could become „more disordered“, which means its entropy may increase.
5. For all $\pi, \lambda \in \Delta^n_+$, it should hold that $(1, \pi) \circ (1, \lambda) = (1, \mathcal{N}(\pi\lambda))$: Towards the limit of maximal confidence, the fusion should correspond to the transition of a Bayesian Prior to the Posterior.

6. For all $\pi, \lambda \in \Delta_+^n$, it should hold that $(0, \pi) \circ (s, \lambda) = (s, \lambda)$: Towards the limit of an update without confidence, the probability-vector should not change.

Now the idea is to construct an operation of the probabilistic semiring on the Bayesian-semiring based on this update function. Therefore consider the following mapping:

$$\alpha :]0,1] \times \mathbb{H} \to \mathbb{H}, \quad \alpha(c, p) := c_\diamond p := p^c, \tag{2}$$

which has the properties (for all $c \in]0,1]$ and arbitrary $p, q \in [0,1]$):

$$\begin{aligned} 2c_\diamond \min\{p,q\} &= \min\{c_\diamond p, c_\diamond q\}, & c_\diamond 0 &= 0, \\ c_\diamond(p \cdot q) &= c_\diamond p \cdot c_\diamond q, & c_\diamond 1 &= 1. \end{aligned}$$

These two equations show that, at least in case $0 < c \leqslant 1$, the mapping $c_\diamond$ is a semiring endomorphism

$$\mathbb{H} \xrightarrow{c_\diamond} \mathbb{H}.$$

Setting additionally

$$\left.\begin{aligned} 20_\diamond : [0,1] \to [0,1], &\quad 0_\diamond p := \begin{cases} 0 & \text{if } p = 0, \\ 1 & \text{if } p > 0, \end{cases} \\ (-\infty)_\diamond : [0,1] \to [0,1], &\quad (-\infty)_\diamond p := \begin{cases} 0 & \text{if } p < 1, \\ 1 & \text{if } p = 1, \end{cases} \end{aligned}\right\} \tag{3}$$

we get, for arbitrary $p, q \in [0,1]$, the properties

$$\begin{aligned} 0_\diamond \min\{p,q\} &= \min\{0_\diamond p, 0_\diamond q\}, \\ 0_\diamond(pq) &= 0_\diamond p \cdot 0_\diamond q, \\ (-\infty)_\diamond \min\{p,q\} &= \min\{(-\infty)_\diamond p, (-\infty)_\diamond q\}, \\ (-\infty)_\diamond(pq) &= (-\infty)_\diamond p \cdot (-\infty)_\diamond q. \end{aligned}$$

Altogether for each $c \in [0,1] \cup \{-\infty\}$ the mapping $c_\diamond : \mathbb{H} \to \mathbb{H}$ is a semiring-endomorphism.

Operation from the Max-Plus-Algebra on the Min-Plus-Algebra

The operation of the probabilistic semiring on the Bayesian semiring given by (2) and (3) can be seen as an operation of the max-plus-algebra on the min-plus-algebra via the above isomorphisms, as in the commutative diagram

$$\begin{array}{ccc} \mathbb{M}\mathrm{ax} \times \mathbb{M}\mathrm{in} & \xrightarrow{\phi} & \mathbb{M}\mathrm{in} \\ {\scriptstyle \mu\times\ell^{-1}}\downarrow & & \uparrow{\scriptstyle \ell} \\ \mathbb{P} \times \mathbb{H} & \xrightarrow[\alpha]{} & \mathbb{H}\,. \end{array}$$

This leads to the mapping

$$\phi : \mathbb{M}\mathrm{ax} \times \mathbb{M}\mathrm{in} \to \mathbb{M}\mathrm{in},$$

$$(a,x) \mapsto \phi(a,x) := a_* x := \begin{cases} \mu(a)x & \text{if } a > 0, \\ 0 & \text{if } a = -\infty \text{ and } x = 0, \\ \infty & \text{if } a = -\infty \text{ and } x > 0; \\ 0 & \text{if } a = 0 \text{ and } x < \infty, \\ \infty & \text{if } a = 0 \text{ and } x = \infty. \end{cases} \tag{4}$$

The following definitions of the products

$$-\infty \cdot 0 = 0, \quad 0 \cdot \infty = \infty, \quad \text{und} \quad \forall x > 0 : \ -\infty \cdot x = \infty,$$

seem unnatural at a first look but are resultig from (3), and, consequently, result from the postulation that $(-\infty)_*$ and 0_* should be endomorphisms on the semiring $\mathbb{M}\mathrm{in}$, whence have to map neutral elements onto neutral elements. Conclusions:

(a) For each $a \in \mathbb{M}\mathrm{ax}$ $a_* : \mathbb{M}\mathrm{in} \to \mathbb{M}\mathrm{in}$ is a semiring endomorphism, which hence particularly keeps neutral elements

$$a_*(\infty) = \infty \quad \text{and} \quad a_* 0 = 0\,. \tag{5}$$

(b) For $0 < a \leq \infty$, we have $0 < \mu(a) \leq 1$, hence $a_* : [0,\infty] \to [0,\infty]$ is bijective. Accordingly for $0 \leq x \leq \infty$ and $0 < a \leq \infty$ the quotient $\frac{x}{\mu(a)}$ is well defined and fulfills

$$(a_*)^{-1} x = \frac{x}{\mu(a)}. \tag{6}$$

The semi-direct product

When building the product of two semirings, one essential trick is first to remove the neutral elemnts before forming the Cartesian product, and re-add them in adequate form afterwards. This leads to the following set:

$$\begin{aligned}\mathbb{K} &:= \Big((\mathbb{M}\mathrm{ax} \setminus \{-\infty, 0\}) \times \mathbb{M}\mathrm{in}\Big) \cup \Big\{(0_{\mathbb{M}\mathrm{ax}}, 0_{\mathbb{M}\mathrm{in}}), (1_{\mathbb{M}\mathrm{ax}}, 1_{\mathbb{M}\mathrm{in}})\Big\} \\ &= \big(]0,\infty] \times [0,\infty]\big) \cup \{(-\infty,\infty),(0,0)\}; \end{aligned} \tag{7}$$

For later usage two properties are noted here, which result from the special role of the two elements $(-\infty, \infty)$ and $(0, 0)$:

$$\left.\begin{array}{rcccl} 2\forall(a,x) \in \mathbb{K}: & a = -\infty & \Rightarrow & x = \infty & \text{and} \quad a_{*}x = \infty, \\ \forall(a,x) \in \mathbb{K}: & a = 0 & \Rightarrow & x = 0 & \text{and} \quad a_{*}x = 0. \end{array}\right\} \tag{8}$$

The $\oplus$-addition over $\mathbb{K}$ will be taking the maximum w.r.t. a total ordering $\leq_{\mathbb{K}}$ on $\mathbb{K}$ formally defined as follows:

$$\forall(a,x),(b,y) \in \mathbb{K}: \quad (a,x) \leq_{\mathbb{K}} (b,y) :\Leftrightarrow \begin{cases} a < b \\ \text{or} \\ a = b \text{ and } x \geq y\,. \end{cases} \tag{9}$$

To motivate this definition, recall that the first component

$$a \in \{-\infty\} \cup [0,\infty]$$

of a confidence-probability pair $(a,x) \in \mathbb{K}$ encodes *confidence*, whereas the second component $x = -\ln p$ is the negative logarithm of a probability. Given two confidence-probability-pairs $(a,x),(b,y) \in \mathbb{K}$, the one with the higher confidence is considered the bigger one. In case of equal confidences, the bigger pair is the one whose second component corresponds to the bigger probability.

Now the $\oplus$-addition over $\mathbb{K}$ is defined by

$$\begin{aligned}(a,x) \oplus (b,y) &:= \max_{\mathbb{K}}\{(a,x),(b,y)\} \\ &:= \begin{cases} (a,x) & \text{if } (b,y) \leq_{\mathbb{K}} (a,x)\,, \\ (b,y) & \text{if } (a,x) \leq_{\mathbb{K}} (b,y)\,. \end{cases}\end{aligned} \tag{10}$$

$\oplus$ is commutative and has the neutral element

$$0_{\mathbb{K}} := (-\infty, \infty),$$

as $(-\infty, \infty)$ is the smallest element of $\mathbb{K}$ w.r.t. the ordering relation $\leq_{\mathbb{K}}$.

To construct the $\otimes$-multiplication we use the following conventions extending the quotient notation (6):

$$\forall x \in \mathbb{M}\text{in} = [0, \infty] : \quad \frac{x}{-\infty} = \infty \quad \text{und} \quad \frac{x}{0} = 0. \tag{11}$$

On a first look, these conventions may seem unnatural, but they have the pleasant consequence

$$\forall (a, x) \in \mathbb{K} : \quad \frac{a_* x}{\mu(a)} = a_* \left(\frac{x}{\mu(a)} \right) = x\,. \tag{12}$$

Indeed, for $0 < a \leq \infty$ the two equations arise from (4). Moreover, for $a = -\infty$ or $a = 0$ we have $\mu(a) = a$, making the equation $\frac{a_* x}{\mu(a)} = x$ a consequence from (8) and the conventions (11), whereas the second equation $a_*(\frac{x}{\mu(a)}) = x$ results from (11) and the homomorphism conditions (5).

The $\otimes$-multiplication over $\mathbb{K}$ is defined through

$$(a, x) \otimes (b, y) := \left(a + b, \frac{a_* x + b_* y}{\mu(a + b)} \right); \tag{13}$$

it is also commutative, and has the neutral element

$$1_{\mathbb{K}} := (0, 0),$$

as is easily seen using (12):

$$\forall (a, x) \in \mathbb{K} :$$
$$(a, x) \otimes (0, 0) = \left(a + 0, \frac{a_* x + 0_* 0}{\mu(a + 0)} \right) = \left(a, \frac{a_* x}{\mu(a)} \right) = (a, x).$$

Associativity und Distributivity

The $\oplus$-addition is just taking the maximum w.r.t. a total ordering, whence associativity of $\oplus$ is immediate.

As on both semirings $\mathbb{M}\mathrm{ax}$ and $\mathbb{M}\mathrm{in}$, the operation $+$ is associative, associativity of the $\otimes$-multiplikation will follow from the assertion

$$\forall (a,x),(b,y),(c,z) \in \mathbb{K}:$$
$$\big((a,x)\otimes(b,y)\big)\otimes(c,z) = \left(a+b+c, \frac{a_*x+b_*y+c_*z}{\mu(a+b+c)}\right),$$

which can be proved as follows:

$$\begin{aligned}\big((a,x)\otimes(b,y)\big)\otimes(c,z) &\overset{(13)}{=} \left(a+b, \frac{a_*x+b_*y}{\mu(a+b)}\right)\otimes(c,z)\\ &\overset{(13)}{=} \left(a+b+c, \frac{(a+b)_*\left(\frac{a_*x+b_*y}{\mu(a+b)}\right)+c_*z}{\mu(a+b+c)}\right)\\ &\overset{(12)}{=} \left(a+b+c, \frac{a_*x+b_*y+c_*z}{\mu(a+b+c)}\right).\end{aligned}$$

Collecting what we've proved up to now, we infer that

$$\big(\mathbb{K},\oplus,(-\infty,\infty),\otimes,(0,0)\big)$$

is a *bimonoid*. In order to prove that this is a semiring, we would have to prove the law of distributivity:

$$\begin{aligned}&\forall (a,x),(b,y),(c,z) \in \mathbb{K}:\\ &\quad\big((a,x)\oplus(b,y)\big)\otimes(c,z) = \big((a,x)\otimes(c,z)\big)\oplus\big((b,y)\otimes(c,z)\big).\end{aligned} \tag{14}$$

Unfortunately, this assertion is false. Counterexamples are provided by those confidence-probability pairs with $a < b$, and $x < y$, $c = \infty$, and $z < \infty$:

$$\begin{aligned}2\big((a,x)\oplus(b,y)\big)\otimes(\infty,z) &= (b,x)\otimes(\infty,z) && \text{because } a<b,\\ &= (\infty, b_*y+z) && \text{by (13), as } \infty_* = \mathrm{id}_{\mathbb{M}\mathrm{in}}.\end{aligned}$$

On the other hand, $a < b$ and $x < y$ imply $a_*x < b_*y$. For $z < \infty$ we infer $a_*x + z < b_*y + z$, and conclude using (10)

$$\begin{aligned}\big((a,x)\otimes(\infty,z)\big)\oplus\big((b,y)\otimes(\infty,z)\big) &= (\infty, a_*x+z)\oplus(\infty, b_*y+z)\\ &= (\infty, a_*x+z),\end{aligned}$$

contradicting (14).

The confidence-probability semiring

With this preliminary work, we are now ready to define the *confidence-probability semiring*. To this end, we remove from $\mathbb{K}$ all pairs with confidence ∞ and define

$$\begin{aligned}\mathbb{K}^\circ &:= \mathbb{K} \setminus \big(\{\infty\} \times \mathbb{M}\mathrm{in}\big)\\ &= \big(]0,\infty[\times[0,\infty]\big) \cup \{(-\infty,\infty),(0,0)\},\end{aligned} \tag{15}$$

and observe that the restriction of (14) to this set is equivalent to the assertion

$$\begin{aligned}&\forall (a,x),(b,y),(c,z) \in \mathbb{K}^\circ :\\ &\quad (a,x) \leq_{\mathbb{K}} (b,y) \quad \Rightarrow \quad (a,x) \otimes (c,z) \leq_{\mathbb{K}} (b,y) \otimes (c,z).\end{aligned}$$

To proof this, assume that $(a,x) \leq_{\mathbb{K}} (b,y)$.

Case 1: $a < b$. Using $c < \infty$ we infer the strict inequality $a + c < b + c$, which implies by (13) and (9) the relation

$$(a,x) \otimes (c,z) \leq_{\mathbb{K}} (b,y) \otimes (c,z).$$

Case 2: $a = b$. Then we conclude from (9) that $x \geq y$, and we have to prove

$$(a,x) \otimes (c,z) \leq_{\mathbb{K}} (a,y) \otimes (c,z),$$

which, by (13), is equivalent to

$$\left(a + c, \frac{a_* x + c_* z}{\mu(a+c}\right) \leq_{\mathbb{K}} \left(a + c, \frac{a_* y + c_* z}{\mu(a+c)}\right). \tag{16}$$

In order to prove this, we have to refer to the second case of (9), and, consequently, we have to prove

$$\frac{a_* x + c_* z}{\mu(a+c)} \geq \frac{a_* y + c_* z}{\mu(a+c)}.$$

If $a + c = 0$ or $a + c = -\infty$, this follows from (11). Otherwise, we have to show

$$a_* x + c_* z \geq a_* y + c_* z. \tag{17}$$

If $z = \infty$, we infer from (5) that $c_* z = \infty$, reducing (17) to the equation $\infty = \infty$.

In case $z < \infty$ it also holds $c_* z < \infty$, which makes (17) equivalent to

$$a_* x \geq a_* y. \tag{18}$$

This follows from $x \geq y$ and the fact that, by their definition in (4), each endomorphism $a_* : \mathbb{M}\text{in} \to \mathbb{M}\text{in}$ is isotonic. ∎

Finally it has been shown that *confidence-probability-semiring*

$$\Big(\mathbb{K}^\circ, \oplus, (-\infty, \infty), \otimes, (0, 0)\Big)$$

has the structure of a commutative semiring. Moreover, the semiring $\mathbb{K}^\circ$ inherits idempotency from the semirings $\mathbb{M}\text{ax}$ and $\mathbb{M}\text{in}$.

Canonical Embedding

The construction of a semidirect product would be incomplete without investigating possible embeddings of the source objects into the product objects. An *embedding* of an object A into an object B is an isomorphism of A onto a subobject of B. Speaking in the terminology of semirings it is an injective homomorphism of semirings of A onto B. In concrete: Let

$$A = (A, \oplus_A, 0_A, \otimes_A, 1_A), \qquad B = (B, \oplus_B, 0_B, \otimes_B, 1_B),$$

be two semirings then an *embedding* of A into B is a mapping $\iota : A \to b$ complying with the following properties:

(i) $\iota(0_A) = 0_B$,

(ii) $\iota(1_A) = 1_B$,

(iii) $\forall x, y \in A : \quad \iota(x) \oplus_B \iota(y) = \iota(x \oplus_A y)$,

(iv) $\forall x, y \in A : \quad \iota(x) \otimes_B \iota(y) = \iota(x \otimes_A y)$,

(v) ι is injective.

In the present case the confidence-probability semiring $\mathbb{K}^\circ$ is seen as the *semidirect product* of the *shortened max-plus-algebra*

$$\mathbb{M}\mathrm{ax}^\circ := \mathbb{M}\mathrm{ax} \setminus \{\infty\}$$

and the min-plus-algebra $\mathbb{M}\mathrm{in}$, written

$$\mathbb{K}^\circ = \mathbb{M}\mathrm{ax}^\circ \ltimes \mathbb{M}\mathrm{in}\,.$$

A homomorphism $\iota : \mathbb{M}\mathrm{ax}^\circ \to \mathbb{K}^\circ$ or $\iota : \mathbb{M}\mathrm{in} \to \mathbb{K}^\circ$ is called a *canonical embedding* if its first resp. second component is the identical mapping, strictly spoken if

$$\forall a \in \mathbb{M}\mathrm{ax}^\circ : \ \iota(a) = (a, \iota_2(a)) \quad \text{resp.} \quad \forall x \in \mathbb{M}\mathrm{in} : \ \iota(x) = (\iota_1(x), x)$$

holds. Obviously a canonical embedding is injective by which the name „embedding“ is reasonable.

Theorem 1 *For every $x \in [0, \infty]$ the mapping*

$$\gamma_x : \mathbb{M}\mathrm{ax}^\circ \to \mathbb{K}^\circ, \quad \gamma_x(a) = \left(a, \frac{a\mu(1) \cdot x}{\mu(a)}\right) = \left(a, \frac{a(1-e^{-1})x}{1-e^{-a}}\right), \qquad (19)$$

is a canonical embedding of the shortened max-plus-algebra into the confidence-probability semiring whereas the conventions from (11) are needed to understand the quotient.

Proof. Let $x \in [0, \infty]$ be fixed. The above mentioned requirements (i)–(v) have to be verified.

At first (i), the preservation of the neutral element of the addition:

$$\gamma_x(-\infty) = \left(-\infty, \frac{-\infty \cdot (1-e^{-1}) \cdot x}{-\infty}\right) = (-\infty, \infty),$$

pursuant to the first part of the conventions from (11).

(ii) concerns the preservation of the neutral element of the multiplication:

$$\gamma_x(0) = \left(0, \frac{0 \cdot (1-e^{-1}) \cdot x}{0}\right) = (0, 0),$$

after the second part of the conventions from (11).

In order to proof (iii), the compatibility of γ_x and the addition, the special character of the second component is irrelevant; so the abbreviation

$$\sigma(a) := \frac{a\mu(1)\cdot x}{\mu(a)}. \tag{20}$$

is used. It holds $\forall a, b \in \mathbb{M}\text{ax}^\circ$:

$$\gamma_x\big(\max\{a,b\}\big) = \begin{cases} \gamma_x(a) = (a, \sigma(a)) & \text{if } a > b, \\ \gamma_x(a) = \gamma_x(b) & \text{if } a = b, \\ \gamma_x(b) = (b, \sigma(b)) & \text{if } a < b, \end{cases}$$
$$= \max_{\mathbb{K}^\circ}\big\{\gamma_x(a), \gamma_x(b)\big\}.$$

For the proof of (iv), the compatibility

$$\forall a, b \in \mathbb{M}\text{ax}^\circ : \qquad \gamma_x(a+b) = \gamma_x(a) \otimes_{\mathbb{K}} \gamma_x(b),$$

the special character of σ matters. As a start for $a = -\infty$ or $b = -\infty$ one gets the equation $a + b = -\infty$ and after (i)

$$\gamma_x(a+b) = \gamma_x(-\infty) = (-\infty, \infty) = 0_{\mathbb{K}}.$$

Let now $a, b \in [0, \infty[$ be arbitrary, then according to (20) and (12)

$$a_*\big(\sigma(a)\big) = a_*\left(\frac{a\mu(1)\cdot x}{\mu(a)}\right) = a\mu(1)\cdot x \tag{21}$$

holds. Now one can begin calculating for $a, b \in [0, \infty[$ according to the proof of (iv):

$$\begin{aligned}
\gamma_x(a+b) &= \big(a, \sigma(a)\big) \otimes_{\mathbb{K}} \big(b, \sigma(b)\big) \\
&\overset{(13)}{=} \left(a+b, \frac{a_*\sigma(a) + b_*\sigma(b)}{\mu(a+b)}\right) \\
&\overset{(21)}{=} \left(a+b, \frac{a\mu(1)\cdot x + b\mu(1)\cdot x}{\mu(a+b)}\right) \\
&= \left(a+b, \frac{(a+b)\mu(1)\cdot x}{\mu(a+b)}\right) \\
&\overset{(21)}{=} (a+b, \sigma(a+b)) \\
&= \gamma_x(a+b).
\end{aligned}$$

The verification of (v) results from the fact that every canonical embedding is per se injective. ■

Theorem 2 *There is a canonical embedding* $\mathbb{M}\text{in} \to \mathbb{K}^\circ$.

Proof. Let $(\iota, \mathrm{id}) : \mathbb{M}\text{in} \to \mathbb{K}^\circ$ be a homomorphism between semirings. In particular

$$\iota(\infty) = -\infty \quad \text{und} \quad \iota(0) = 0.$$

hold. Because of (8) it follows:

$$\forall x \in \mathbb{M}\text{in} \setminus \{0, \infty\} \quad a := \iota(x) > 0. \tag{22}$$

Let $x \in \mathbb{M}\text{in} \setminus \{0, \infty\}$ be arbitrary. On the one hand now

$$(\iota, \mathrm{id})(x + x) = (\iota(x + x), x + x), \tag{23}$$

holds and on the other hand, because (ι, id) is a homomorphism between semirings, it also holds after (iii)

$$\begin{aligned}(\iota, \mathrm{id})(x + x) &= (\iota, \mathrm{id})(x) \otimes_{\mathbb{K}} (\iota(x), x) = (a, x) \otimes_{\mathbb{K}} (a, x) \\ &= \left(a + a, \frac{a_* x + a_* x}{\mu(a + a)}\right) = \left(a + a, \frac{\mu(a)(x + x)}{\mu(a + a)}\right).\end{aligned}$$

If $0 < \mu(a) < 1$, one could continue with positiv reals and would get because of (23)

$$1 = \frac{\mu(a)}{\mu(a + a)} = \frac{1 - e^{-a}}{1 - e^{-2a}}.$$

From that would follow $2a = a$. Because of $a \in \mathbb{M}\text{ax}^\circ = \{-\infty\} \cup [0, \infty[$, this would imply either $a = 0$ or $a = -\infty$, in contradiction to assumption $0 < \mu(a) < 1$. Hence $\iota(x) = a \in \{-\infty, 0\}$ and by (8)

$$\forall x \in \mathbb{M}\text{in} : \ (\iota(x), x) \in \{(-\infty, \infty,), (0, 0)\},$$

but that is impossible after the construction of $\mathbb{K}^\circ$ from (7). ■

Finally, let us note that the mapping

$$\iota : \mathbb{M}\text{in} \to \mathbb{K}, \quad \iota(x) := \begin{cases} (0, 0) & \text{if } x = 0, \\ (-\infty, \infty) & \text{if } x = \infty, \\ (\infty, x) & \text{otherwise}, \end{cases}$$

is a canonical embedding of the min-plus-algebra into the bimonoid $\mathbb{K}$.

References

[1] Markus Huber. *Semantische Informationsbearbeitung unter Berücksichtigung von Konfidenz-Zahlen*. Diplomarbeit in Mathematik. Katholische Universität Eichstätt-Ingolstadt, 2009.

[2] Mehryar Mohri. Weighted Automata Algorithms. In Manfred Droste, Werner Kuich, and Heiko Vogler, editors, *Handbook of Weighted Automata*, Monographs in Theoretical Computer Science, pages 213–254. 2009.