Limit of generalized belief fusion operator

Sergey Kirgizov, Nicolas Gastineau, and Lobna Azaza

LE2I UMR6306, CNRS, Arts et Métiers, *Univ. Bourgogne Franche-Comté, F-21000 Dijon, France* `firstname.lastname@u-bourgogne.fr`

Abstract

We consider a generalized belief combination operator in which the set intersection is replaced by a symmetric binary operator. Using classical Markov chain theory we study the convergence of the sequence of such combinations. We show how to apply our results to the case of classical belief combination operator.

Keywords: Markov chain applications, Dempster-Shafer theory, combination operator, convergence

MCS2010: 60J10, 60J20, 91D30

1 Introduction: motivation and definitions

The belief functions theory is considered as a framework for reasoning under uncertainty, it is connected to other similar frameworks such as probability, possibility and imprecise probability theories. This theory has applications in various fields: for example, in recent years there is increasing interest from researchers who study social [2], neural [3] and other complex networks [10].

As far as we are concerned, the question of the convergence of the iterative application of combination rules [1], has not been considered yet, despite its importance for contemporain works. For example, in [1] authors associate different belief masses for each Twitter relationship type: one mass for retweet, another for mention and yet another mass for reply. Using these masses they estimate the influence/popularity of a Twitter user using the combination of the belief masses of his mentions, retweets of his tweets, and replies to him. The cumulative number of retweets, replies and mentions only grows over time. Even if we restrict ourselves to some specific period of time this number can be very large (hundreds of thousands) for some Twitter accounts. In this paper, using Markov chain theory, we study the convergence of the iterative application of a generalized combination rule. The idea of generalized combination rule goes back to works [5, 11]. In this paper, we use a generalized rule defined in [1].

In Subsection 1.1 we recall basic notions from classical belief theory. In Subsection 1.2 we describe our generalization and formulate our main question about the conditions under which the sequence of iterative applications of the combination rule converges. In order to answer this question we use classical Markov chains theory. For the sake of consistency, we recall main notions and theorems from this theory (Subsection 1.3). In Section 2 we discuss properties of the generalized combination rule (also known as *Dempster-Shafer-Smets combination operator*), we show whether the iterative application of a combination rule converges and how our results can be applied to the classical Dempster-Shafer-Smets theory. Finally, we discuss possible directions of the further research in Section 3.

[1] We use the phrases "combination operator" and "combination rule" interchangeable.

1.1 Dempster–Shafer–Smets operator

The theory of belief functions, also known as evidence theory or Dempster-Shafer theory, was firstly introduced by Dempster in the context of statistical inference, and was later developed by Shafer and Smets [7, 9].

In the following, we are going to remind the basic concepts of belief functions theory. Let Ω be a finite set, denote by $\mathcal{P}(\Omega)$ the set of all subsets of Ω. A *mass* m is a function $m : \mathcal{P}(\Omega) \to [0, 1]$ such that $\sum_{X \in \mathcal{P}(\Omega)} m(X) = 1$. The mass $m(X)$ expresses the part of belief that supports the subset X of Ω.

Belief functions theory allows, not only the representation of the partial knowledge, but also the information fusion [8]. This is usually done by the conjunctive combination rule [9]. Considering two mass functions m_1 and m_2, the conjunctive combination rule, denoted by $\textcircled{\cap}$, is defined as follows:

$$(m_1 \textcircled{\cap} m_2)(C) = \sum_{\substack{A \cap B = C \\ A \in \mathcal{P}(\Omega) \\ B \in \mathcal{P}(\Omega)}} m_1(A) m_2(B), \quad \forall C \in \mathcal{P}(\Omega) \qquad \text{(classical combination rule)}$$

1.2 Generalized belief functions and the question of convergence

In classical Dempster-Shafer-Smets theory we should use $\mathcal{P}(\Omega)$ as a domain of mass functions. In this paper we propose to use any finite set Λ in place of $\mathcal{P}(\Omega)$. In this generalized setting, a mass function have the following type $m : \Lambda \to [0, 1]$ and satisfies the equation $\sum\limits_{x \in \Lambda} m(X) = 1$.

The *generalized combination rule* (we use the symbol $\otimes$) is a modified version of classical combination rule: instead of using the set intersection operator we consider any symmetric operator $@ : \Lambda^2 \to \Lambda$.

$$(m \otimes m')(z) = \sum_{\substack{x \,@\, y = z \\ x \in \Lambda \\ y \in \Lambda}} m(x) m'(y), \quad \forall z \in \Lambda \qquad \text{(modified combination rule)}$$

Note that $\otimes$ depends on m, m' and $@$. This rule was defined in this form for the first time by Azaza *et al.* in [1]. A similar modification of the classical rule was proposed by Ben Dhaou *et al.* [2]. Let $m^{\otimes n} = \underbrace{\Big(\big((m \otimes m) \otimes m\big) \otimes \cdots \otimes m\Big)}_{n \text{ times}}$. Let $\otimes$ be left associative, so we can omit parentheses in what follows. Consider the point-wise limit $\lim_{n \to \infty} m^{\otimes n} : \Lambda \to [0, 1]$ defined as:

$$\lim_{n \to \infty} m^{\otimes n} = K \iff \exists K : \Lambda \to [0, 1] \text{ s.t. } \forall x \in \Lambda \lim_{n \to \infty} m^{\otimes n}(x) = K(x)$$

In this paper we show how the existence of this limit depends on m and $@$.

1.3 Markov chains

We recall basic notions from the Markov chain theory. Additional information and proofs can be found in Chapter 4 of [6].

Let I be a *set of states*. An *initial distribution* $\tau : I \to [0, 1]$ is a probability distribution defined over the set of states, that is $\sum_{x \in I} \tau(x) = 1$. Denote by $x \to y$ a *transition* from state x to state y. Note that $(x \to y) \in I^2$. Here we allow ourselves to use a bit of aesthetic sugar and write $x \circlearrowleft$ in place of $x \to x$. A *transition distribution* $T : I^2 \to [0, 1]$ is a probability distribution defined over all transitions, such that for all $x \in I$ we have $\sum_{y \in I} T(x \to y) = 1$. A *Markov chain* is a triplet (I, τ, T). A Markov process can be described inductively:

Base case. The process starts at the state x with the probability $\tau(x)$.

Inductive step. The process proceeds from the state y to the state z with the probability $T(y \to z)$.

It is useful to think about the transition distribution as a right stochastic (the sum of each row is equal to 1) matrix T. We define $T_{i,j} = T(i \to j)$ and represent the initial distribution τ by a vector. Using the matrix notation, the Markov process described above, can be written as τT^n. The study of convergence of the Markov chain naturally translates into the study of $\lim_{n\to\infty} \tau T^n$.

Denote by $x \rightsquigarrow y$ a sequence of transitions with strictly positive probabilities that starts at x and ends at y. We call this sequence of transitions a *path*. In general there are several paths from x to y, if we want to distinguish them we use the following notion: $x \rightsquigarrow_1 y$, $x \rightsquigarrow_2 y$, etc. In order to denote the set of all paths we use $P_{x \rightsquigarrow y}$.

Denote by $|p|$ the number of transitions in a path p. The *period* $d(x)$ of a state x is defined as $d(x) = \gcd(\{|p| : p \in P_{x \rightsquigarrow x}\})$. A state x is said to be *periodic* if $d(x) > 1$. A state x is said to be *aperiodic* if $d(x) = 1$. Sometimes there are no paths from x to x, in this case $d(x) = \gcd(\varnothing)$ is undefined, abusing the notation, we also say that such state x is aperiodic. A Markov chain is said to be aperiodic when all its states are aperiodic.

An *absorbing state* t, denoted by $\boxed{\mathrm{t}}$, is a state with no outgoing paths, except the self-loop $\boxed{\mathrm{t}} \circlearrowleft$. All absorbing states are aperiodic, but the converse is not true in general.

When we have $x \rightsquigarrow y$ and $y \rightsquigarrow x$, we say that x and y *communicate*. Also, we say that x communicates with itself. A *communication class* is a maximal subset of states $J \subseteq I$ such that any pair of states from J communicates. In this way the set of all states is paritioned into several mutually disjoint subsets (communication classes) $I = \bigcup_{i=1}^{k} J_i$, where k is the number of communication classes. We denote by $J[x]$ the communication class of x.

A Markov chain is said to be *irreducible* if there is only one communication class, i.e all pairs of states $x, y \in I$ communicate. Otherwise, a Markov chain is said to be *reducible* and can be reduced into several irreducible communication classes. It is useful to represent the structure of a reducible Markov chain as a directed acyclic graph, where nodes are the communication classes, and where there is a directed edge $J_i \to J_k$ between two different communication classes J_i and J_k if and only if there is a transition with strictly positive probability $x \to y$ for some $x \in J_i$ and $y \in J_k$. A communication class that has no outgoing edges is called a *sink class*. When a sink class contains only one state, this state is absorbing. If we leave a class, we cannot return back. Thus, after a sufficiently large number of transitions only sink classes will matter.

A Markov chain with a transition distribution T and an initial distribution τ is said to be *convergent* when $\lim_{n\to\infty} \tau T^n$ exists. When this limit exists and does not depend on τ we say that the Markov chain has *unique limiting distribution*.

A *stationary distribution* π is a vector such that $\pi T = \pi$. The unique limiting distribution is a stationary one. Any finite Markov chains has a stationary distribution. But, some Markov chains do not have the unique limiting distribution.

Classical results about convergence of finite Markov chains [6] can be summarised in the following manner:

Finite irreducible aperiodic	$\Longrightarrow$	Converges. Limit *does not depend* on the initial distribution.
Finite aperiodic	$\Longrightarrow$	Converges. Limit *depends* on the initial distribution. Only absorbing states may have a positive masses at the limit.
All states are periodic	$\Longrightarrow$	Does not converge *in general* (but converges when the initial distribution is a stationary one).
Only some states are periodic	$\Longrightarrow$	We should examine the structure of the directed acyclic graph of communication classes and see what happens in sink classes: some sink classes may be periodic, other may not; depending on initial distribution some sink classes may have null mass at the limit.

2 Properties of generalized combination rule and the question of convergence

Results discussed in this subsection are not so surprising, but again for the sake of consistency, we need to state them clearly. We start this section by discussing some properties of the generalized combination rule $\otimes$ and the symmetric operation @ that replaces the intersection operator in the classical rule. Next, we answer the question about the existence of $\lim_{n\to\infty} m^{\otimes n}$. Finally, we show how our results can be applied to the study of convergence of classical Dempster-Shafer-Smets operator.

2.1 Properties of $\otimes$

Proposition 2.1. *A combination of two mass functions is another mass function.*

Proof. Denote $(m \otimes m')$ by m''. It is easy to see that for all x we have $m''(x) \geqslant 0$, because we compute m'' using only multiplication and addition of a non-negative numbers. Next, we show that $\sum_{z\in\Lambda} m''(z) = 1$. Let $\Lambda_z^2 = \{(x,y) \in \Lambda : x @ y = z\}$ and proceed as follows:

$$\sum_{z\in\Lambda} m''(z) = \sum_{z\in\Lambda} \sum_{\substack{x @ y = z \\ x\in\Lambda \\ y\in\Lambda}} m(x)m'(y)$$
$$= \sum_{z\in\Lambda} \sum_{(x,y)\in\Lambda_z^2} m(x)m'(y)\,.$$

Note that $\Lambda_z^2 \neq \Lambda_{z'}^2 \iff z \neq z'$, and $\bigcup_{z\in\Lambda} \Lambda_z^2 = \Lambda^2$. So, we can omit $\sum_{z\in\Lambda}$ and rewrite as follows:

$$= \sum_{(x,y)\in\Lambda^2} m(x)m'(y)$$
$$= \sum_{x\in\Lambda} \sum_{y\in\Lambda} m(x)m'(y)$$
$$= \sum_{x\in\Lambda} m(x) \sum_{y\in\Lambda} m'(y)$$

m and m' are mass functions: $\sum_{x\in\Lambda} m(x) = \sum_{y\in\Lambda} m'(y) = 1$, so

$$\sum_{x\in\Lambda} m(x) \sum_{y\in\Lambda} m'(y) = 1$$

□

From the definition of $\otimes$ and symmetry of @ it follows that $\otimes$ is symmetric, that is $m \otimes m' = m' \otimes m$.

Proposition 2.2. *In general $\otimes$ is non-associative, i.e. there exist mass functions m and m' such that the following equation does not hold*

$$(m \otimes m') \otimes m'' \neq m \otimes (m' \otimes m'')$$

Proof. Consider $\Lambda = \{A, B, C\}$, and the following @:

@	A	B	C
A	B	B	C
B	B	C	C
C	C	C	C

Let

$$m = m' = \begin{array}{|c|c|c}A & B & C \\ \hline 1 & 0 & 0\end{array}$$

$$m'' = \begin{array}{|c|c|c}A & B & C \\ \hline 0 & 1 & 0\end{array}$$

It is easy to see that:

$$(m \otimes m') \otimes m'' = \begin{array}{|c|c|c}A & B & C \\ \hline 0 & 0 & 1\end{array}$$

$$m \otimes (m' \otimes m'') = \begin{array}{|c|c|c}A & B & C \\ \hline 0 & 1 & 0\end{array}$$

Thus, in general:

$$(m \otimes m') \otimes m'' \neq m \otimes (m' \otimes m'')$$

□

Also, the following equation does not hold in general

$$(m \otimes m) \otimes (m \otimes m) = \big((m \otimes m) \otimes m\big) \otimes m$$

Let's try

$$m = \begin{array}{c|c|c|c} & A & B & C \\ \hline m & 1 & 0 & 0\end{array}$$

$$m \otimes m = m^{\otimes 2} = \begin{array}{c|c|c|c} & A & B & C \\ \hline m & 0 & 1 & 0\end{array}$$

$$m \otimes m \otimes m = m^{\otimes 2} \otimes m = \begin{array}{c|c|c|c} & A & B & C \\ \hline m & 0 & 1 & 0\end{array}$$

$$m \otimes m \otimes m \otimes m = m^{\otimes 3} \otimes m = \begin{array}{c|c|c|c} & A & B & C \\ \hline m & 0 & 1 & 0\end{array}$$

$$(m \otimes m) \otimes (m \otimes m) = m^{\otimes 2} \otimes m^{\otimes 2} = \begin{array}{c|c|c|c} & A & B & C \\ \hline m & 0 & 0 & 1\end{array}$$

As a consequence we have $m^{\otimes 2n} \neq m^{\otimes n} \otimes m^{\otimes n}$ thus the order of combinations may change drastically the result.

2.2 From @ to the Markov chain

Using @, $\otimes$ and a mass function m we construct a Markov chain, denoted by $M_{@}$, as follows:

- The elements of Λ are the states of the chain.
- m is the initial distribution.
- The transition probability $T(x \to z)$ is defined as $T(x \to z) = \sum_{x\,@\,y=z, y\in\Lambda} m(y)$. For example, $x\,@\,y = z$ and $x\,@\,y' = z$ correspond to the part of the Markov chain presented in Figure 1. Note that $\sum_{z\in\Lambda} T(x \to z) = \sum_{x\,@\,y=z;\, y,z\in\Lambda} m(y) = 1$.

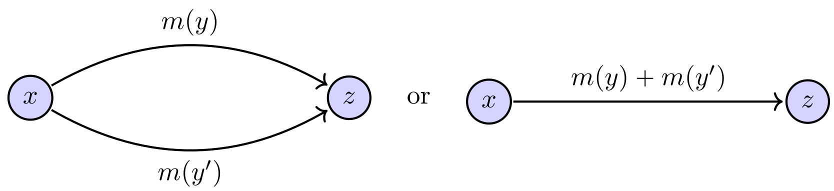


Figure 1 – A part of a Markov chain that appears when we have $x\,@\,y = z$ and $x\,@\,y' = z$.

Observation 2.3. *The existence of* $x \xrightarrow{m(y)} z$ *implies the existence of* $y \xrightarrow{m(x)} z$ *because* @ *is symmetric. In a special case of* $x = y$ *these two transitions coincide and we have only one transition* $x \xrightarrow{m(x)} z$.

Recall that $\otimes$ is defined to be left associative, and observe that

$$m \underbrace{\overbrace{\otimes m}^{T} \overbrace{\otimes m}^{T} \cdots \overbrace{\otimes m}^{T}}_{\text{n times}} = mT^n$$

Thus, any question about $\lim_{n\to\infty} m^{\otimes n}$ naturally translates into a question about $\lim_{n\to\infty} mT^n$.

Example. Consider $\Lambda = \{A, B, C, D, E\}$, and let @ be defined as follows:

	A	B	C	D	E
A	B	C	D	D	A
B	C	C	D	D	B
C	D	D	D	D	C
D	D	D	D	D	D
E	A	B	C	D	E

Figure 2a represents the transition matrix constructed from this @. A partial construction of a Markov chain is shown at Figure 2b. The full Markov chain can be seen at Figure 2c.

@

	A	B	C	D	E
A	B	C	D	D	A
B	C	C	D	D	B
C	D	D	D	D	C
D	D	D	D	D	D
E	A	B	C	D	E

$\Longrightarrow$

T

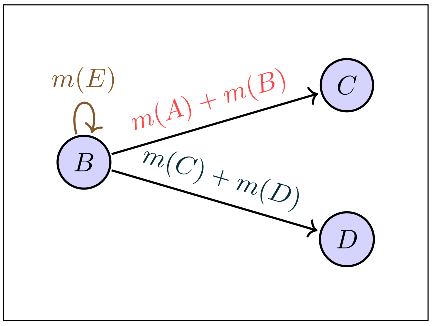
	A	B	C	D	E
A	$m(E)$	$m(A)$	$m(B)$	$m(C)+m(D)$	0
B	0	$m(E)$	$m(A)+m(B)$	$m(C)+m(D)$	0
C	0	0	$m(E)$	$m(A)+m(B)+m(C)+m(D)$	0
D	0	0	0	$m(B)+m(C)+m(D)+m(E)$	0
E	$m(A)$	$m(B)$	$m(C)$	$m(D)$	$m(E)$

(a) Transition matrix constructed from @

@

	A	B	C	D	E
A	B	C	D	D	A
B	C	C	D	D	B
C	D	D	D	D	C
D	D	D	D	D	D
E	A	B	C	D	E

$\Longrightarrow$

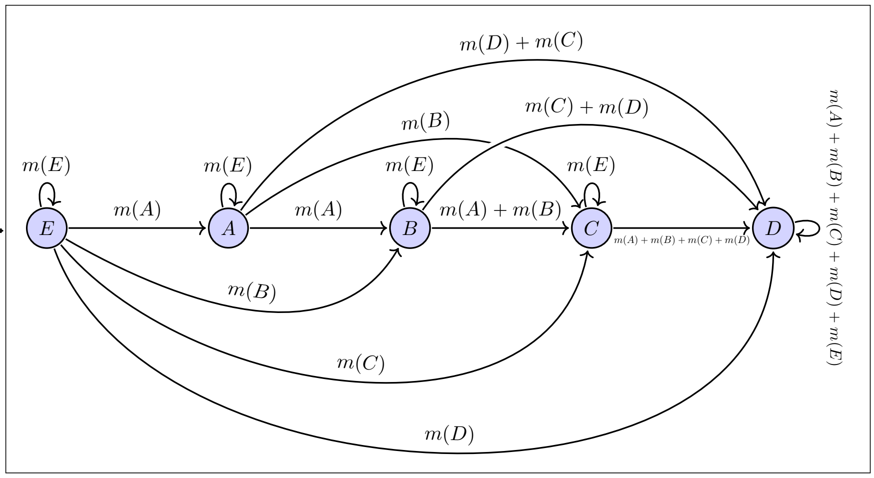


(b) Partial view of a Markov chain

@

	A	B	C	D	E
A	B	C	D	D	A
B	C	C	D	D	B
C	D	D	D	D	C
D	D	D	D	D	D
E	A	B	C	D	E

$\Longrightarrow$



(c) Full view of a Markov chain

Figure 2 – A Markov chain constructed from @

2.3 Properties of the constructed Markov chain

Using $@ : \Lambda^2 \to \Lambda$ we define a transitive binary relation $\preceq$ over Λ as follows:

$$\begin{aligned} &x@y = z \quad \Rightarrow \quad x \lhd z \text{ and } y \lhd z \\ &\text{Let } \preceq \text{ be a transitive closure of } \lhd \end{aligned} \tag{$\preceq$}$$

Observation 2.4. *Suppose that all masses are positive, i.e.* $\forall x \in \Lambda : m(x) > 0$*. Consider a pair of states* $x, y \in \Lambda$*. It is not difficult to see that there is a path* $x \rightsquigarrow y$ *if and only if* $x \preceq y$*.*

Markov chain is said to be *weakly connected* when its underlying undirected graph is connected. Two states x and y are weakly connected, if there is a path between them in the underlying undirected graph.

Proposition 2.5. *When all masses are positive, the Markov chain* $M_@$ *is weakly connected.*

Proof. Consider any pair of states $x, y \in \Lambda$. Recall that $@ : \Lambda^2 \to \Lambda$, so there exists $z \in \Lambda$ such that $x@y = z$. By the definition of $\preceq$, we have $x \preceq z$ and $y \preceq z$. Using Observation 2.4 we obtain $x \rightsquigarrow z$ and $y \rightsquigarrow z$. Hence, x and y are weakly connected via z. □

The situation is slightly more complicated if some masses are null. Consider, for example, the Markov chain from Figure 2c: let $m(E) = 1$, and see that there are no paths between different states. In order to study the cases when only some masses are positive we proceed as follows:

Let $A(x)$ be the set of accessible states from the state x, i.e. $A(x) = \{y \in \Lambda \text{ s.t. } \exists x \rightsquigarrow y\} \cup \{x\}$. Let $\mathcal{A}(\Upsilon) = \bigcup_{x\in\Upsilon} A(x)$. Denote by $M_@^\Upsilon$ the Markov chain created from $M_@$ by removing all states that are not in $\mathcal{A}(\Upsilon)$ together with all corresponding transitions (i.e. we remove a transition $x \to y$ if $x \notin \mathcal{A}(\Upsilon)$ or $y \notin \mathcal{A}(\Upsilon)$), and by setting $m(x) = 0$ for all $x \in \Lambda \backslash \Upsilon$. This subchain arises when all states from $\Lambda \backslash \Upsilon$ have null mass.

Before proving that $M_@^\Upsilon$ subchain is weakly connected for all $\Upsilon \subseteq \Lambda$, we prove the following lemma.

Lemma 2.6. *Any pair of states* $x, y \in \Lambda$ *with positive masses is weakly connected.*

Proof. We have $x@y = z$ for some $z \in \Lambda$. By construction of the Markov chain we have $x \xrightarrow{m(y)} z \xleftarrow{m(x)} y$. We also have $m(x) > 0$ and $m(y) > 0$, so there are $x \rightsquigarrow z$ and $y \rightsquigarrow z$. Therefore, x and y are weakly connected. □

Proposition 2.7. *For all* $\Upsilon \subseteq \Lambda$ *the subchain* $M_@^\Upsilon$ *is weakly connected.*

Proof. Consider any pair of different states x, y from $M_@^\Upsilon$. Four possibilities exist:

- When $m(x) > 0$ and $m(y) > 0$ we use Lemma 2.6 to show that x and y are weakly connected.
- When $m(x) > 0$ and $m(y) = 0$ there are two subcases:
  - If $y \in A(x)$, we have $x \rightsquigarrow y$.
  - Otherwise, there is a state $z \neq x$ with a positive mass, such that $y \in A(z)$ and $z \in \Upsilon$, since the Markov chain $M_@^\Upsilon$ contains only the states from $\bigcup_{u\in\Upsilon} A(u)$. So, we have $z \rightsquigarrow y$ and (by Lemma 2.6) x and z are weakly connected. Hence, x and y are weakly connected.
- The case when $m(y) > 0$ and $m(x) = 0$ is symmetrically equivalent to the previous one.

- When $m(x) = m(y) = 0$ we have two subcases:
  - There is a state z with positive mass such that $x, y \in A(z)$. In this case, we have $z \rightsquigarrow x$ and $z \rightsquigarrow y$. Thus, x and y are weakly connected.
  - There are two different states z and w states with positive masses such that $x \in A(z)$ and $y \in A(w)$. So, we have $z \rightsquigarrow x$, $w \rightsquigarrow y$. By Lemma 2.6) z and w are weakly connected. Hence, x and y are weakly connected.

□

2.3.1 Non-necessary reflexive poset of states

By definition, $\preceq$ is a transitive binary relation defined on Λ. Sometimes this relation is antisymmetric, i.e. if $x \preceq y$ and $y \preceq x$ then $x = y$. A set together with an antisymmetric transitive binary relation is called the *non-necessary reflexive poset*: for some x the relation is reflexive, i.e. $x \preceq x$, but for another x it is irreflexive, i.e. $x \npreceq x$.

Proposition 2.8. *When $(\Lambda, \preceq)$ is a non-necessary reflexive poset, there is unique maximal element in this poset.*

Proof. Suppose that there are two different maximal elements x and y. Recall that $@ : \Lambda^2 \to \Lambda$, so there is $z \in \Lambda$ such that $x@y = z$. By the definition of $\preceq$ we have $x \preceq z$. Since, element x is maximal, so $z = x$, but in this case we have $x@y = x$ and $y \preceq x$, thus y cannot be maximal. □

Proposition 2.9. *When $(\Lambda, \preceq)$ is a non-necessary reflexive poset, a Markov chain $M_@$ always converges.*

Proof. Recall that $x \preceq z$, obtained from $x@y = z$, corresponds to $x \xrightarrow{m(y)} z$ in the Markov chain. When $(\Lambda, \preceq)$ is a non-necessary reflexive poset, there is no cycles in the Markov chain besides self-loops, because if there is a cycle $x \to z \to \ldots \to x$ then we should have

1. $x \preceq z \preceq \ldots \preceq x$ (by transitivity)
2. $x = z$ (by antisymmetry)

Thus, any state in the Markov chain is aperiodic. Finite aperiodic chains always converge. □

When $(\Lambda, \preceq)$ is a non-necessary reflexive poset, the limit $\lim_{n\to\infty} \tau T^n$ may depend on the initial distribution τ, but only absorbing (i.e. with self-loops and without outgoing transitions of positive mass) states may have a positive masses at the limit. Consider a state x with a self-loop $x \circlearrowleft_{m(x)}$. Let $m(x) = 1$: all mass will rests at the state x forever. If there is only one state t with a self-loop $t \circlearrowleft_{m(t)}$, then we have $(\lim_{n\to\infty} \tau T^n)(t) = 1$ for any initial distribution τ.

The Markov chain from Figure 2c can have only two convergent states E and D. Moreover, $(\lim_{n\to\infty} \tau T^n)(E) = 1$ if and only if $m(E) = 1$.

2.3.2 Strict poset of communication classes

Denote by $\mathcal{J}$ the set of all communication classes. Recall that a Markov chain can be represented as a directed acyclic graph (DAG), where nodes are the communication classes, and where there is a directed edge $J_i \to J_k$ between two different communication classes J_i and J_k if and only if there is a transition wtih strictly positive probability $x \to y$ for some $x \in J_i$ and $y \in J_k$. Suppose that all masses are positive, and define a binary relation $<$ over the set of all communication classes $\mathcal{J}$ in natural way:

$$\begin{array}{c} J_i \to J_k \quad \Rightarrow \quad J_i \lessdot J_k \\ \text{Let } < \text{ be a transitive closure of } \lessdot \end{array} \tag{<}$$

The set of all communication classes $\mathcal{J}$ together with $<$ form a strict (irreflexive) poset.

Proposition 2.10. *When all state masses are positive, i.e.* $\forall x \in \Lambda : m(x) > 0$*, the DAG of communication classes of a Markov chain* $M_@$ *is weakly connected.*

Proof. The DAG of communication classes is created by fusion of some states, so the weak connectivity of the DAG follows from the weak connectivity of the Markov chain (see Proposition 2.5). □

Proposition 2.11. *For all* $\Upsilon \subseteq \Lambda$ *the DAG of communication classes of a Markov chain* $M_@^\Upsilon$ *is weakly connected.*

Proof. This proof is identical to the previous proof, but we use Proposition 2.7 instead of Proposition 2.5. □

Lemma 2.12. *Consider two states* x *and* y*, if there is a path* $x \rightsquigarrow y$ *then we have only two possibilities:* $J[x] = J[y]$ *or* $J[x] < J[y]$*.*

Proof. There is a path $x \rightsquigarrow y$ thus, if $J[x] \neq J[y]$, we have $J[x] < J[y]$ by the definition of the DAG of communication classes. □

Proposition 2.13. *When all state masses are positive, the DAG of communication classes of a Markov chain* $M_@$ *have only one sink class. Equivalently: the poset* $(\mathcal{J}, <)$ *have only one maximal element.*

Proof. Suppose there are two maximal elements J and J'. Take some $x \in J$ and $y \in J'$. Consider $x@y = z$. By construction of $M_@$ we have $x \rightsquigarrow z$. So, by Lemma 2.12 we obtain $J[x] = J[z]$ or $J[x] < J[z]$. In the latter case we have a contradiction (as we supposed that J is maximal). Symmetrically, from $y@x = z$ we obtain $J[y] = J[z]$. Hence, $J[y] = J[z] = J[x]$, that means $J = J'$. □

If there are some states with null masses, the DAG of communication classes of the Markov chain constructed from Λ and @ may have several sinks. Consider Figure 3, and let $m(X) = 0.5$, $m(Y) = 0.5$. In this case, the state Z will have the mass 0.5 at the limit, and, in addition, there are two periodic communication classes: $\{W, W'\}$ and $\{V, V'\}$. It is easy to see (by looking only at thick black transitions), that this Markov chain have three sinks: Z, $\{W, W'\}$ and $\{V, V'\}$.

The following theorem and its corollaries follow naturally from our propositions:

Theorem 2.14. $\lim_{n\to\infty} m^{\otimes n}$ *exists if and only if all sink classes of the Markov chain* $M_@^\Upsilon$ *are aperiodic, where* $\Upsilon = \{x \in \Lambda \text{ s.t. } m(x) > 0\}$*. Only the states from sink classes may have positive masses at the limit.*

By the Proposition 2.13, the DAG of communication classes of a Markov chain $M_@^\Upsilon = M_@$ contains only one sink class when all masses are positive. Thus, we obtain the following corollary.

Corollary 2.15. *When all masses are positive, i.e.* $\Upsilon = \Lambda$*,* $\lim_{n\to\infty} m^{\otimes n}$ *exists if and only if the unique sink is aperiodic.*

@

	X	Y	Z	W	W'	V	V'
X	W	Z	Z	W'	W	V'	V
Y	Z	V	Z	W'	W	V'	V
Z	Z	Z					
W	W'	W'					
W'	W	W			Z		
V	V'	V'					
V'	V	V					

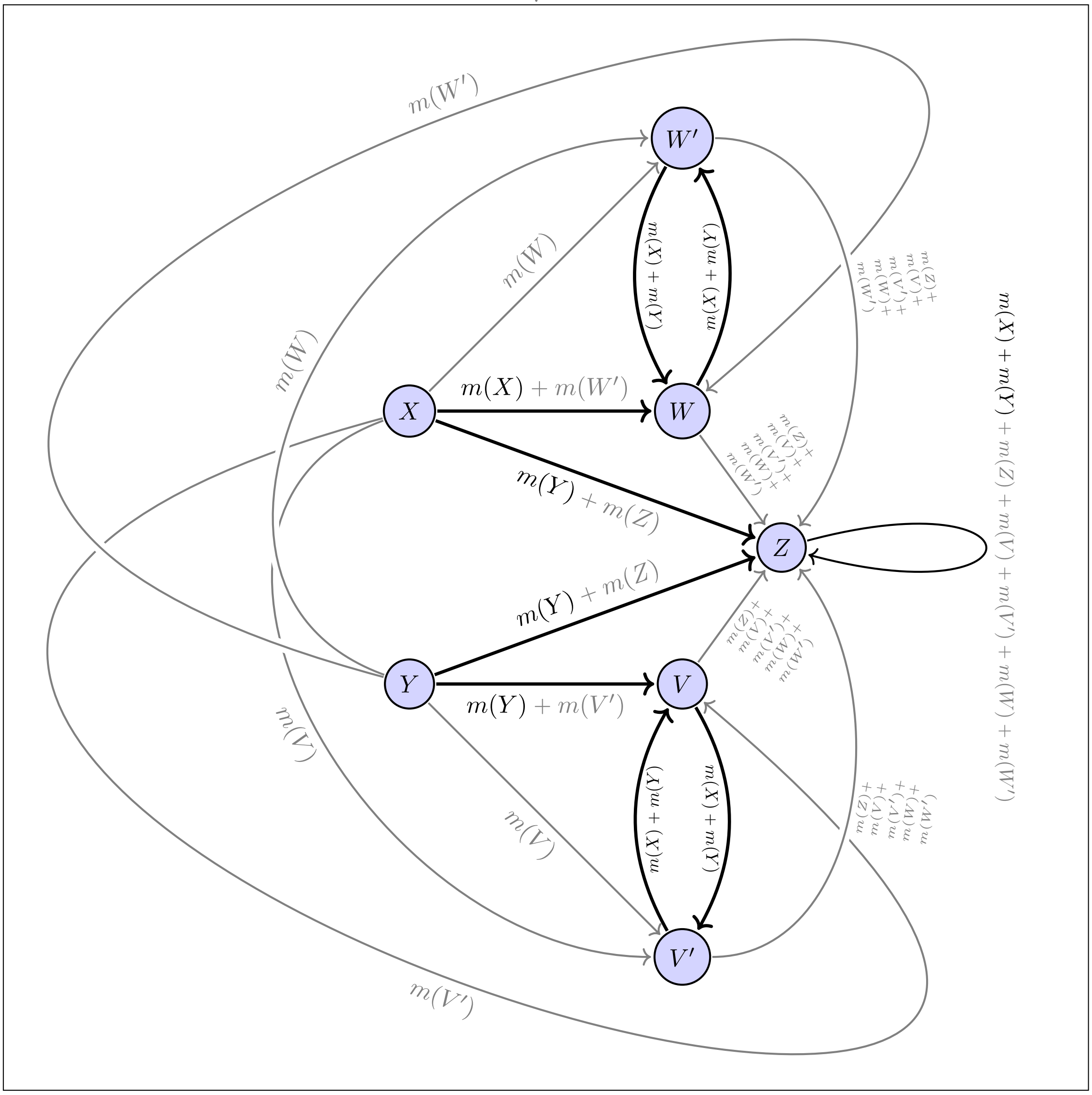


Figure 3 – A butterfly-like Markov chain constructed from @. When all states have positive masses, Z is the unique sink. But if only $m(X) > 0$ and $m(Y) > 0$, there are three sinks: $Z, \{W, W'\}, \{V, V'\}$.

Corollary 2.16. *The following statements are equivalent:*

- $\lim_{n\to\infty} m^{\otimes n}$ *exists for all* m*.*
- *For all* $\Upsilon \subseteq \Lambda$ *the DAG of communication classes of a Markov chain* $M_{@}^{\Upsilon}$ *has only aperiodic sink classes.*

2.4 Convergence of classical Dempster-Shafer-Smets operator

In the classical case (see Subsection 1.1) $\Lambda = \mathcal{P}(\Omega)$ and $@$ is the set intersection operator $\cap$. The equation $X \cap Y = Z$ translates into $X \preceq Z$ and $Y \preceq Z$. Observe that $\preceq$ in this case behaves precisely like the set inclusion operator $\subseteq$. The set $\mathcal{P}(\Omega)$ of subsets of Ω ordered by $\subseteq$ is a classical example of a poset so, by Proposition 2.9, $\lim_{n\to\infty} \underbrace{m \circledcirc m \circledcirc \cdots \circledcirc m}_{n \text{ times}}$ always exists. Denote this limit by m^*. If only the states from $\Upsilon \subseteq \mathcal{P}(\Omega)$ have positive masses, then the limit converges to its intersection, i.e. to the state $s = \bigcap_{x\in\Upsilon} x$. Thus, we have $m^*(s) = 1$ at the limit. In particular, if all masses are positive, the only sink is the empty set, and we have $m^*(\varnothing) = 1$. Denote by $M_{\circledcirc}^{\Upsilon}$ the Markov chain associated with the iterative application of the classical Dempster-Shafer-Smets combination rule when only the states from Υ have positive masses. Figure 4 illustrates the Markov chain $M_{\circledcirc}^{\Upsilon}$, where $\Upsilon = \big\{\{A,B,C\},\{A,B\},\{A,C\},D\big\}$.

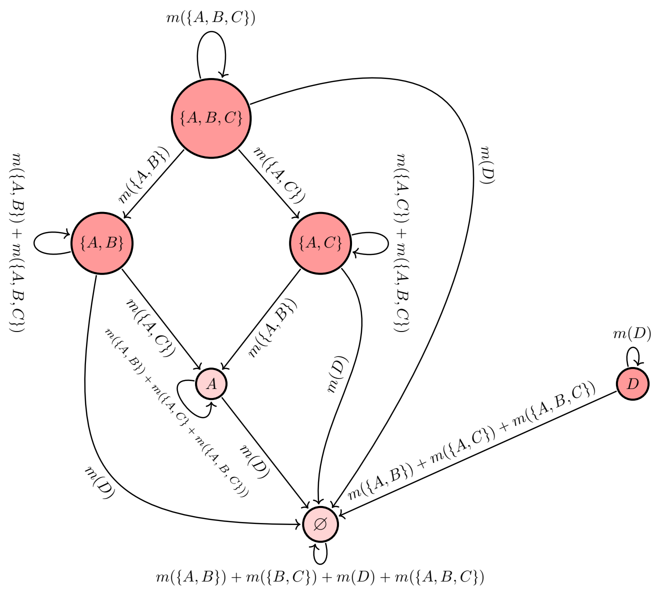


Figure 4 – A Markov chain associated with a classical Dempster-Shafer-Smets combination rule. $\Omega = \{A, B, C, D\}$. Initially all masses are null except $m(\{A, B, C\})$, $m(\{A, B\})$, $m(\{A, C\})$ and $m(D)$. The whole mass "flows" to the state $\varnothing$. In the case of the normalised classical combination rule, only the states A and D may have a positive mass at the limit.

2.5 Convergence of normalised Dempster-Shafer-Smets operator

Here we consider the *normalised version of the classical rule*, that is

$$\begin{aligned}(m_1 \boxed{\cap} m_2)(\varnothing) &= 0\\ (m_1 \boxed{\cap} m_2)(C) &= \tfrac{(m_1 \circledcirc m_2)(C)}{(m_1 \circledcirc m_2)(\varnothing)} \quad \forall C \in \mathcal{P}(\Omega)\backslash\{\varnothing\}\end{aligned} \qquad \text{(normalised classical rule)}$$

Denote by `MAXS` the set of sinks of the Markov chain obtained from $M^{\Upsilon}_{@}$ by removing $\{\varnothing\}$. For example, in Markov chain from Figure 4 we have $\texttt{MAXS} = \{A, D\}$.

Let $m_1 = m$, $m_n = m_{n-1} \boxed{\cap} m_1$, and consider the behaviour of $\lim_{n\to\infty} m_n$.

Proposition 2.17. *In the case of the iterative application of the normalised classical Dempster-Shafer-Smets combination rule, the whole mass will be (at the limit) distributed between elements of* `MAXS`*.*

Proof. Recall that $x \preceq y$, obtained from $x \cap y = z$, corresponds to $x \xrightarrow{m(y)} z$ in the Markov chain $M^{\Upsilon}_{\circledcirc}$. Take any minimal (in $\preceq$-sense) state $Y \notin \texttt{MAXS}$ with a positive mass. Note that $m(Y) \neq 1$, otherwise it should be in `MAXS`. Thus, there exists the following chain of subsets $Y_k \subset Y_{k-1} \subset Y_{k-2} \subset \cdots \subset Y_1 = Y$ such that for any $i \in [1, k-1]$ there is a transition of positive mass $Y_i \to Y_{i+1}$ and $Y_k \in \texttt{MAXS}$ (for example, in Markov chain from Figure 4 we have $\{A\} \subset \{A, B\} \subset \{A, B, C\}$).

In what follows, we compare $m_n(Y_2)$ and $m_n(Y_1)$, and we show that $\big(\lim_{n\to\infty} m_n\big)(Y_1) = 0$. Next, we use the induction in order to show that only the "last" state $Y_k \in MAXS$ may have a positive mass at the limit.

Let $X = Y_2$. X is connected to Y by a transition $Y \to X$ of a positive mass. Let $m(Y \to X)$ be the mass of the transition from Y to X. The only incoming transition to Y is a self-loop $Y \circlearrowleft_{m_1(Y)}$, thus we have

$$(m_{n-1} \circledcirc m_1)(Y) = m_{n-1}(Y) \cdot m_1(Y)$$

The situation is a bit more complex around the state X, where we have a self-loop, a transition $Y \to X$, and, possibly, other states $Z_1, Z_2, \ldots, Z_\ell$ with transitions $Z_i \to X$ for any $i \in [1, \ell]$:

$$\begin{tikzcd} Y \arrow[r, "m(Y\to X)"] & X \arrow[loop above, "m_1(Y)+m(X)+\epsilon'"] & Z_i \arrow[l, "\epsilon_i"'] \end{tikzcd} \qquad i \in [1, \ell]$$

So, we can write

$$(m_{n-1} \circledcirc m_1)(X) = m_{n-1}(X) \cdot \big(m_1(Y) + m(Y \to X) + \epsilon'\big) + \epsilon$$

where, $\epsilon = \sum_{i=1}^{\ell} \epsilon_i$ and ϵ' are some non-negative masses. We expect the ratio between $m_n(Y)$ and $m_n(X)$ as follows:

$$\frac{m_n(Y)}{m_n(X)} = \frac{\frac{(m_{n-1} \circledcirc m_1)(Y)}{(m_{n-1} \circledcirc m_1)(\varnothing)}}{\frac{(m_{n-1} \circledcirc m_1)(X)}{(m_{n-1} \circledcirc m_1)(\varnothing)}} = \frac{(m_{n-1} \circledcirc m_1)(Y)}{(m_{n-1} \circledcirc m_1)(X)} \leqslant$$

$$\frac{m_{n-1}(Y) \cdot m_1(Y)}{m_{n-1}(X) \cdot \big(m_1(Y) + m(Y \to X)\big)} \leqslant K \frac{m_{n-1}(Y)}{m_{n-1}(X)}$$

where $K = \frac{m_1(Y)}{m_1(Y) + m_1(Y \to X)} < 1$. We obtain $\frac{m_n(Y)}{m_n(X)} \leqslant \frac{m_1(Y)}{m_1(X)} K^n$, so $\lim_{n\to\infty} \frac{m_n(Y)}{m_n(X)} = 0$, and, in particular, $\lim_{n\to\infty} m_n(Y) = 0$. □

3 Open questions

From practical points of view it may be important to consider other combinations rules (see for exemple [4]), improve the Proposition 2.17 by considering the question *"How exactly the whole mass is distributed over* `MAXS`*?"*, discuss convergence speed and numerical methods of limit calculation, answer the question *"When a transition matrix constructed from symmetric binary operator is diagonalisable or not"*. The following theoretical question seems to be very intriguing: given a Markov chain M, is there some @ such that $M_{@}$ is isomorphic to M? Another interesting direction consists in the study of typical (randomly generated) symmetric binary operation defined over the set Λ, when $|\Lambda| \to \infty$, and when the cardinality of Λ exceeds $\aleph_0$.

Acknowledgements

Authors are grateful to Éric Leclercq, Mahmoud Omidvar, and Armen Petrossian for helpful comments and numerous discussions. This work was partly funded by a TEE 2014 project, PEPS MOMIS SEISME project, and Burgundy region.

Bibliography

References

[1] L. Azaza, S. Kirgizov, M. Savonnet, E. Leclercq, and R. Faiz, *Influence assessment in twitter multi-relational network*, in Signal Image Technologies and Internet-Based System (SITIS), Eleventh International IEEE Conference on, IEEE, 2015, pp. 436–443.

[2] S. Ben Dhaou, M. Kharoune, A. Martin, and B. Ben Yaghlane, *Belief approach for social networks*, in Belief Functions: Theory and Applications, Springer, 2014, pp. 115–123.

[3] T. Denoeux, *A neural network classifier based on dempster-shafer theory*, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 30 (2000), pp. 131–150.

[4] T. Denœux, *Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence*, Artificial Intelligence, 172 (2008), pp. 234–264.

[5] D. Dubois and H. Prade, *A set-theoretic view of belief functions logical operations and approximations by fuzzy sets*, International Journal Of General System, 12 (1986), pp. 193–226.

[6] R. G. Gallager, *Stochastic processes: theory for applications*, Cambridge University Press, 2013.

[7] S. Kotz and W. N. L. Johnson eds., *Belief functions*, Encyclopedia of Statistical Sciences 1 209, (1982).

[8] V. Nimier and A. Appriou, *Utilisation de la théorie de dempster-shafer pour la fusion d'informations*, GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, (1995), pp. 137–140.

[9] P. Smets, *Imperfect Information: Imprecision and Uncertainty*, in Uncertainty Management in Information Systems, A. Motro and P. Smets, eds., Springer US, 1997, pp. 225–254.

[10] D. Wei, X. Deng, X. Zhang, Y. Deng, and S. Mahadevan, *Identifying influential nodes in weighted networks based on evidence theory*, Physica A: Statistical Mechanics and its Applications, 392 (2013), pp. 2564–2575.

[11] R. R. Yager, *Arithmetic and other operations on dempster-shafer structures*, International Journal of Man-Machine Studies, 25 (1986), pp. 357–366.