Classic Probability Revisited (II): Algebraic Operations of the Extended Probability Theory

YINGXU WANG
International Institute of Cognitive Informatics and Cognitive Computing (ICIC)
Laboratory for Computational Intelligence, Software Science, and Denotational Mathematics
Dept. of Electrical and Computer Engineering
Schulich School of Engineering and Hotchkiss Brain Institute
University of Calgary
2500 University Drive, NW, Calgary, Alberta, CANADA T2N 1N4

Abstract: - Part II of this paper presents a set of comprehensive algebraic operators on the extended mathematical structures of the general probability theory. It is recognized that the classic probability theory is cyclically defined among a small set of highly coupled operations. In order to solve this fundamental problem, a reductive framework of the general probability theory is introduced. It is found that conditional probability operation on consecutive events is the key to independently manipulate other probability operations. This leads to a revisited framework of rigorous manipulations on general probabilities. It also provides a proof for a revisited Bayes' law fitting in more general contexts of variant sample spaces and complex event relations in fundamental probability theories. The revisited probability theory enables a rigorous treatment of uncertainty events and causations in formal inference, qualification, quantification, and semantic analysis in contemporary fields such as cognitive informatics, computational intelligence, cognitive robots, complex systems, soft computing, and brain informatics.

Key-Words: - Denotational mathematics, probability theory, probability algebra, fuzzy probability, formal inference, cognitive informatics, cognitive computing, computational intelligence, semantic computing, brain informatics, cognitive systems

1. Introduction

Probability theory is a branch of mathematics that deals with uncertainty and probabilistic norms of random events and potential causations as well as their algebraic manipulations. The development of classic theories of probability can be traced back to the work of Blaise Pascal (1623-1662) and Pierre de Fermat (1601-1665) [Todhunter, 1865; Venn, 1888; Hacking, 1975]. Many others such as Jacob Bernoulli, Reverend T. Bayes, and Joseph Lagrange had significantly contributed to probability theory. Theories of probability in its modern form was unified by Pierre Simon and Marquis de Laplace in the 19th century [Kolmogorov, 1933; Whitworth, 1959; Hacking, 1975; Mosteller, 1987; Bender, 1996]. Set theories [Cantor, 1874; Zadeh, 1965, 1968, 1996, 2002; Artin, 1991; Ross, 1995; Pedrycz & Gomide, 1998; Novak et al., 1999; Potter, 2004; Gowers, 2008; BISC, 2013; Wang, 2007] provide an expressive power for modeling the discourse and axioms of probability theories. A theory of fuzzy probability and its algebraic framework has been presented in [Wang, 2015e].

The philosophy of probability theory is analogy-based where large-enough experiments are required for establishing prior probability estimations and norms in a certain sample space. The main methodology of classic probability theory is an external or black box predication for a set of uncertain phenomena of a complex system without probing into its internal mechanisms. Although the range of prior probability for any predicated event is [0, 1], the range of posterior probability is merely reduced to {0, 1} immediately after the given event has realized in a certain probability space.

It is recognized that the classic probability theory is cyclically defined among a set of highly coupled operations where only logically conjunctive, disjunctive, and conditional events are considered. This paper presents a revisited theory of probability, which extends classic probability theory to a comprehensive set of probability operations. An extended set of algebraic operators on the revisited mathematical model of probability is rigorously defined in Section 2, which extends the traditional probability operations of addition, multiplication, and condition to subtraction and division. The conventional mutual-coupled probability operations are independently separated in a deductive structure on the basis of the refined model of conditional probability. Formal properties of probability and rules of algebraic operations on general probabilities are summarized in Section 3. Proven theorems and practical examples are provided throughout the paper for elaborating each of the fundamental definitions and operations in the general theory of probability. The revisited theory of probability may be used to solve a number of challenging problems in classic probability theory such as complex sequential, concurrent, and causal probabilities as well as real-time probabilities under highly restrictive timing constraints.

Due to its excessive length, this paper is presented in two parts on: i) The mathematical models of general probability; and ii) The algebraic operations of the extended probability theory. This paper is the second part of the general probability theory on formal algebraic operators on the extended mathematical structures of the general probability theory.

2. Algebraic Operations on the General Probability Models

The theoretical framework of general probability is formally represented by the mathematical model and the set of algebraic operators of probability as outlined in Sections 3 and 4 of Part I. A set of six probability operators are identified as those of conditional, multiplication, division, addition, subtraction, and complement operations, which stem from a unified mathematical model of the conditional probability. Each probability operator is formally defined and elaborated in the following subsections towards an algebraic framework of the theory of general probability.

2.1 The Conditional Operator on Consecutive Probabilities

The conditional operation of consecutive probabilities deals with coupled influences between related events in both invariant and variant sample spaces. Because conditional probability forms the foundation for all other operators in the algebraic system of the general probability theory, it must be rigorously analyzed in order to avoid the dilemma of the cyclic definition as in classic probability theory.

The nature of conditional probability is constrained by different contexts determined by three control factors of the Cartesian product, $S \times R \times D$, as defined in Table 1 where S demotes the sample space (variant/invariant), R relation of events (joint/disjoint), and D dependency of events (dependent/ independent/mutually-exclusive (ME)). Therefore, the contexts of the general probability are classified into four categories according to the probability characteristics in the Cartesian product, i.e.: a) *invariant* sample space and *disjoint/*ME-*dependent* events, b) *invariant* sample space and *joint/independent* events, c) *variant* sample space and *disjoint/independent* events, and d) *invariant* sample space and *joint/dependent* events.

Table 1. Contexts of Relations and Dependencies of Events in the General Probability Theory

No	Category	Definition ($S \times R \times D$)	Sample space (S)	Events	
				Relation (R)	**Dependency (D)**
i	Disjoint/mutually-exclusive (ME) events in invariant sample space	$S \times \overline{R} \times D$	$S = S'$	$X \cap Y = \varnothing$	$X \to (Y = \varnothing)$, ME
ii	Joint/independent events in invariant sample space	$S \times R \times \overline{D}$		$X \cap Y \neq \varnothing$	$X \not\to Y$
iii	Disjoint/independent events in variant sample space	$S' \times \overline{R} \times \overline{D}$	$S \neq S'$	$X \cap Y = \varnothing$	$X \not\to Y$
iv	Joint/dependent events in variant sample space	$S' \times R \times D$		$X \cap Y \neq \varnothing$	$X \to (Y = Y')$

Theorem 1. The *conditional operator on consecutive probabilities* of an event b influenced by that of a preceding event a in the sample space S in $\mathfrak{U}$, $P(b \mid a)$, is determined by a ratio between the changed sizes of sets of succeeding events B' and of the sample space S' given $a \in A \subset S$, and $b \in B \subset S'$, i.e.:

$$P(b \mid a) \triangleq P(A \to B)$$
$$= \begin{cases} i) \text{ Invariant } S, \text{unrelated } \overline{R}, \text{ and ME-dependent } D\text{: } S \times \overline{R} \times D \\ \quad 0 \\ ii) \text{ Invariant } S, \text{ related } R, \text{ and independent } \overline{D}\text{: } S \times R \times \overline{D} \\ \quad P(b) \\ ii) \text{ Variant } S', \text{ unrelated } \overline{R}, \text{ and independent } \overline{D}\text{: } S' \times \overline{R} \times \overline{D} \\ \quad P'(b) = \dfrac{P(b)}{1 - P(a_i)},\ P(a_i) = \dfrac{P(a)}{|A|} \\ iv) \text{ Variant } S', \text{ related } R \text{ and dependent } D\text{: } S' \times R \times D \\ \quad P''(b) = \dfrac{P(b) - P(b_i)}{1 - P(b_i)},\ P(b_i) = \dfrac{P(b)}{|B|} \end{cases} \tag{1}$$

Proof. Theorem 1 can be proven in each of the four contexts as defined in Table 1 according to Definition 10 in Part I as follows:

$$\forall a, b, a \in A \subset S, \text{ and } b \in B \subset S',$$
$$P(b \mid a) = P(A \to B) = \frac{|B'|}{|S \setminus a|}, a \in A \wedge b \in B$$
$$= \begin{cases} i) \ \ \forall S' = S \wedge A \cap B = \varnothing \wedge A \to (B = \varnothing, \text{ME}), \\ \quad \frac{|\varnothing|}{|S|} = 0 \\ ii) \ \ \forall S' = S \wedge A \cap B \neq \varnothing \wedge A \not\to B, \\ \quad \frac{|B|}{|S|} = P(b) \\ iii) \ \forall S' \neq S \wedge A \cap B = \varnothing \wedge A \not\to B, \\ \quad \frac{|B|}{|S|-1} = \frac{P(b)}{1 - P(a_i)} = P'(b), \ P(a_i) = \frac{P(a)}{|A|} \\ iv) \ \forall S \neq S \wedge A \cap B \neq \varnothing \wedge A \to B', \\ \quad \frac{|B|-1}{|S|-1} = \frac{P(b) - P(b_i)}{1 - P(b_i)} = P''(b), \ P(b_i) = \frac{P(b)}{|B|} \end{cases} \quad (2)$$

■

It is noteworthy in Theorem 1 that *S* is variant in general as constrained by Theorem 1 in Part I because of the coupling of the conditional events *a* and *b*. In other words, the general probability in an invariant sample space is only a special case of that of the general variant context.

Example 1. In an *invariant* sample space $S_1 = \{H = 0.68, \ T = 0.32\}$ as modeled in Example 2 in Part I, the events *head* (*H*) and *tail* (*T*) are mutually exclusive in a single toss of the coin. That is, both events cannot happen simultaneously. Once a head is observed, tail will certainly not appear in the same trail, and vice versa. This is a typical context of *mutually exclusive* ($H \cap T = \varnothing$ or *disjoint*), and *dependent* ($T \mid H = \varnothing$ or $H \mid T = \varnothing$) events of conditional probability according to Theorem 1(i) where $P(T \mid H) = 0, \ if \ H \cap T = \varnothing \text{ and } H \to (T = \varnothing)$.

It is noteworthy that a pair of mutually exclusive events X and Y are dependent because $X \cap Y = \varnothing \wedge X \to (Y = \varnothing) \Rightarrow P(Y \mid X) = 0$, due to the interactive influence between the non-independent events.

Example 2. Given a bag containing five black balls (*B*) and five white balls (*W*) in $S_2 = \{\mathop{R}_{i=1}^{5} P(b_i \mid b_i \in B) = 0.11, \ \mathop{R}_{i=6}^{10} P(w_i \mid w_i \in W) = 0.09\}$ as modeled in Example 3 in Part I. Assume the ball drawn from the bag will be returned to the bag before the next trial, i.e., $S_2' = S_2$, it is a case of invariant sample space, related and independent events of conditional probability according to Theorem 1(i) as follows:

$$P(W \mid B) = P(W) = 0.45$$
$$P(B \mid W) = P(B) = 0.55$$
$$P(W \mid W) = P(W) = 0.45$$
$$P(B \mid B) = P(B) = 0.55$$

Example 3. Reconsider Example 2 in $S_2' = \{\mathop{R}_{i=1}^{5} P(b_i \mid b_i \in B) = 0.11, \ \mathop{R}_{i=6}^{10} P(w_i \mid w_i \in W) = 0.09\}$ where the ball drawn from the bag will not be returned, i.e., $S_2' \neq S_2$, it becomes a case of *variant* sample space, disjoint /independent or joint/dependent events of conditional probability according to Theorem 1(iii) or 1(iv), respectively, as follows:

$$P(W \mid B) = P'(W) = \frac{P(W)}{1 - P(b_i)}, \ P(b_i) = \frac{P(B)}{|B|} = 0.55 / 5 = 0.11$$
$$= \frac{0.45}{1 - 0.11} = \frac{0.45}{0.89} = 0.51$$
$$P(B \mid W) = P'(B) = \frac{P(B)}{1 - P(w_i)}, \ P(w_i) = \frac{P(W)}{|W|} = 0.45 / 5 = 0.09$$
$$= \frac{0.55}{1 - 0.09} = \frac{0.55}{0.91} = 0.60$$
$$P(B \mid B) = P''(B) = \frac{P(B) - P(b_i)}{1 - P(b_i)}, \ P(b_i) = 0.11$$
$$= \frac{0.55 - 0.11}{1 - 0.11} = \frac{0.44}{0.89} = 0.49$$
$$P(W \mid W) = P''(W) = \frac{P(W) - P(w_i)}{1 - P(w_i)}, \ P(w_i) = 0.09$$
$$= \frac{0.45 - 0.09}{1 - 0.09} = \frac{0.36}{0.91} = 0.40$$

Contrasting the results obtained in Examples 2 and 3, it is noteworthy that the conditional probabilities in Contexts (iii) and (iv) of Theorem 1 have increased or decreased, respectively, due to the size shrinkages of sample spaces and/or the number of events as a result of the conditional coupling. The changes between the variant (S_2') and invariant (S_2) sample space can be analyzed as follows:

$$\begin{cases} P'(W \mid B) - P(W \mid B) = 0.51 - 0.45 = 0.06 \\ P'(B \mid B) - P(B \mid B) = 0.49 - 0.55 = -0.06 \\ P'(B \mid W) - P(B \mid W) = 0.60 - 0.55 = 0.05 \\ P'(W \mid W) - P(W \mid W) = 0.40 - 0.45 = -0.05 \end{cases}$$

The results indicate that conditional probabilities in the variant and invariant sample spaces may be significantly different due to the increment or decrement of coupled influences.

2.2 The Complement Operator on the Context of Probability

Theorem 2. The *complement of probability of an event* $a \in A \subset S$ in $\mathfrak{U}$, $P(\overline{a})$, is determined by the probability of all events in *S* excluding only that of *a*, i.e.:

$$P(\overline{a}) \triangleq 1 - P(a) \quad (3)$$

Proof. Theorem 2 can be proven according to Definition 10 in Part I as follows:

$$\begin{aligned}&\forall a \in A \subset S \text{ and } \bar{a} \in \bar{A} \subset S,\\ &P(\bar{a}) = P(\bar{A} \,|\, a \in A \subset S \wedge \bar{a} \in \bar{A} \subset S)\\ &\quad = \frac{|\bar{A}|}{|S|} = \frac{|S \setminus A|}{|S|} = \frac{|S|}{|S|} - \frac{|A|}{|S|}\\ &\quad = 1 - P(a)\end{aligned} \tag{4}$$

■

Example 4. On the basis of Example 4 in Part I, the complement of probability in the sample space $S_1^2 = \{HH = 0.46,\ HT = 0.22,\ TH = 0.22,\ TT = 0.10\}$ can be determined according to Theorem 2 as follows:

$$\begin{aligned}P(\overline{HH}) &= 1 - P(HH) = 1 - 0.46 = 0.54\\ P(\overline{TH}) &= 1 - P(TH) = 1 - 0.22 = 0.78\end{aligned}$$

Corollary 1. The *double complements of the general probability* of an event $a \in A \subset S$ in $\mathfrak{U}, P(\bar{\bar{a}})$, results in an involution to the same probability, i.e.:

$$P(\bar{\bar{a}}) = 1 - P(\bar{a}) = 1 - (1 - P(a)) = P(a) \tag{5}$$

2.3 The Multiplication Operator on Disjunctive Probabilities

Theorem 3. The *multiplication of probabilities* of disjunctive events *a* and *b* in the sample space *S* in $\mathfrak{U}, P(a \times b)$, is determined by the product of the probabilities of $P(a)$ and $P(b \,|\, a)$ given $a \in A \subset S$ and $b \in B \subset S'$, i.e.:

$$P(a \times b) \triangleq P(A \wedge B) = P(a)P(b \,|\, a)$$
$$= \begin{cases} i) \ \text{Invariant } S, \text{unrelated } \bar{R}, \text{ and ME-dependent } D\text{: } S \times \bar{R} \times D \\ \quad 0 \\ ii) \ \text{Invariant } S, \text{ related } R, \text{ and independent } \bar{D}\text{: } S \times R \times \bar{D} \\ \quad P(a)P(b) \\ iii) \ \text{Variant } S', \text{ unrelated } \bar{R}, \text{ and independent } \bar{D}\text{: } S' \times \bar{R} \times \bar{D} \\ \quad P(a)P'(b) = P(a)\dfrac{P(b)}{1 - P(a_i)},\ \ P(a_i) = \dfrac{P(a)}{|A|} \\ iv) \ \text{Variant } S', \text{ related } R \text{ and dependent } D\text{: } S' \times R \times D \\ \quad P(a)P''(b) = P(a)\dfrac{P(b) - P(b_i)}{1 - P(b_i)},\ \ P(b_i) = \dfrac{P(b)}{|B|} \end{cases} \tag{6}$$

Proof. Theorem 3 can be proven according to Definition 10 in Part I and Theorem 1 as follows:

$$\begin{aligned}&\forall a, b, a \in A \subset S, \text{ and } \mathrm{b} \in B \subset S',\\ &P(a \times b) = P(A \wedge B) = \frac{|A \cap B|}{|S|} = \frac{|A|}{|S|} \bullet \frac{|B \setminus a_i|}{|S \setminus a_i|},\ a_i \in A\\ &\quad = P(a)P(b \,|\, a)\end{aligned}$$

$$= \begin{cases} i) \ \ \forall S' = S \wedge A \cap B = \varnothing \wedge A \to (B = \varnothing, \text{ME}), \\ \quad 0 \\ ii) \ \ \forall S' = S \wedge A \cap B \neq \varnothing \wedge A \not\to B, \\ \quad P(a)P(b) \\ iii) \ \forall S' \neq S \wedge A \cap B = \varnothing \wedge A \not\to B, \\ \quad P(a)\dfrac{P(b)}{1 - P(a_i)} = P(a)P'(b),\ \ P(a_i) = \dfrac{P(a)}{|A|} \\ iv) \ \forall S \neq S \wedge A \cap B \neq \varnothing \wedge A \to B', \\ \quad P(a)\dfrac{P(b) - P(b_i)}{1 - P(b_i)} = P(a)P''(b),\ P(b_i) = \dfrac{P(b)}{|B|} \end{cases} \tag{7}$$

■

Example 5. Given an *invariant* sample space $S_1 = \{H = 0.68,\ T = 0.32\}$ as modeled in Example 2 in Part I, i.e., $S_1^{'} = S_1$, the following disjunctive probabilities for two consecutive tosses of the uneven coin can be derived by a probability multiplication according to Theorem 3(ii):

$$\begin{aligned}P(H \times T) &= P(H)P(T) = 0.68 \bullet 0.32 = 0.22\\ P(H \times H) &= P(H)P(H) = 0.68 \bullet 0.68 = 0.46\\ P(T \times H) &= P(T)P(H) = 0.32 \bullet 0.68 = 0.22\\ P(T \times T) &= P(T)P(T) = 0.32 \bullet 0.32 = 0.10\end{aligned}$$

Example 6. Given a *variant* sample space $S_2 = \{\mathop{R}\limits_{i=1}^{5} P(b_i \,|\, b_i \in B) = 0.11,\ \mathop{R}\limits_{i=6}^{10} P(w_i \,|\, w_i \in W) = 0.09\}$ as modeled in Example 3 in Part I, i.e., $S_2^{'} \neq S_2$, the following probability multiplications for two consecutive draws of the uneven balls in the bag can be obtained according to Theorem 3(iii) or 3(iv), respectively:

$$\begin{aligned}P(B \times W) &= P(B)P'(W) = \frac{P(B)P(W)}{1 - P(w_i)},\ P(w_i) = 0.09\\ &= \frac{0.68 \bullet 0.32}{1 - 0.09} = \frac{0.22}{0.91} = 0.24\\ P(W \times B) &= P(W)P'(B) = \frac{P(W)P(B)}{1 - P(b_i)},\ P(b_i) = 0.11\\ &= \frac{0.32 \bullet 0.68}{1 - 0.11} = \frac{0.22}{0.89} = 0.25\\ P(B \times B) &= P(B)P''(B) = P(B)\frac{P(B) - P(b_i)}{1 - P(b_i)},\ P(b_i) = 0.11\\ &= \frac{0.68(0.68 - 0.11)}{1 - 0.11} = \frac{0.39}{0.89} = 0.44\\ P(W \times W) &= P(W)P''(W) = P(W)\frac{P(W) - P(w_i)}{1 - P(bw)},\ P(w_i) = 0.09\\ &= \frac{0.32(0.32 - 0.09)}{1 - 0.09} = \frac{0.07}{0.91} = 0.08\end{aligned}$$

Corollary 2. The *revisited Bayes' law of probability* can be rigorously derived based on Theorem 3 as follows:

$$\begin{aligned}
&\forall a,b, a \in A \subset S,\ b \in B \subset S',\ \text{and } S' = S,\\
&P(a \times b) = P(a)P(b \mid a),\ A \cap B \neq \varnothing \wedge A \not\to B\\
&\qquad = P(a)P(b)\\
&\qquad = P(b)P(a \mid b)\\
&\qquad = P(b \times a)\\
&\Rightarrow \begin{cases} \dfrac{P(b \mid a)}{P(b)} = \dfrac{P(a \mid b)}{P(a)}, & S' = S \wedge A \cap B \neq \varnothing \wedge A \not\to B \wedge B \not\to A \\ \dfrac{P(b \mid a)}{P(b)} \neq \dfrac{P(a \mid b)}{P(a)}, & \text{Otherwise} \end{cases}
\end{aligned} \tag{8}$$

Corollary 2 and Theorem 3 indicate that Bayes' law in classic probability theory is a special case of general probability multiplication, which may only hold *iff* $S' = S \wedge A \cap B \neq \varnothing \wedge B \mid A = B$, i.e., when the conditions for invariant sample space and related but independent events are satisfied.

2.4 The Division Operator on Composite Probabilities

The algebraic operation of probability division is an inverse operation of probability multiplication, which is not defined in traditional probability theory.

Theorem 4. The *division of probability* of an event *b* by that of another event *a* in the sample space *S* in $\mathfrak{U}$, $P(b/a)$, is determined by the ratio of their probabilities where $a \in A \subset S$ and $b \in B \subset S'$ i.e.:

$$P(b/a) \triangleq P(\frac{|B|}{|A|}) = \frac{P(b)}{P(a)},\ 0 < P(a) \geq P(b)$$

$$= \begin{cases} i) \text{ Invariant } S, \text{unrelated } \overline{R}, \text{ and ME dependent } D\text{: } S \times \overline{R} \times D \\ \quad 0 \\ ii) \text{ Invariant } S, \text{ related } R, \text{ and independent } \overline{D}\text{: } S \times R \times \overline{D} \\ \quad \dfrac{P(b)}{P(a)} \\ iii) \text{ Variant } S', \text{ unrelated } \overline{R}, \text{ and independent } \overline{D}\text{: } S' \times \overline{R} \times \overline{D} \\ \quad \dfrac{P'(b)}{P(a)} \\ iv) \text{ Variant } S', \text{ related } R \text{ and dependent } D\text{: } S' \times R \times D \\ \quad \dfrac{P''(b)}{P(a)} \end{cases} \tag{9}$$

Proof. Theorem 4 can be proven according to Definition 10 in Part I as well as Theorems 1 and Theorem 3 as follows:

$$\begin{aligned}
&\forall a,b, a \in A \subset S,\ \text{and } b \in B \subset S',\\
&P(b/a) = P(|B|/|A|) = \frac{|B|}{|A|}\\
&\qquad = \frac{|B|/|S|}{|A|/|S|},\ 0 < |A| \geq |B|\\
&\qquad = \frac{P(b)}{P(a)},\ 0 < P(a) \geq P(b)\\
&= \begin{cases} i)\ \forall S' = S \wedge A \cap B = \varnothing \wedge A \to (B = \varnothing, \text{ME}),\\ \quad \dfrac{|B|/|S|}{|A|/|S|} = \dfrac{P(b)}{P(a)} = \dfrac{0}{P(a)} = 0,\ A \to (B = \varnothing) \\ ii)\ \forall S' = S \wedge A \cap B \neq \varnothing \wedge A \not\to B,\\ \quad \dfrac{|B|/|S|}{|A|/|S|} = \dfrac{P(b)}{P(a)} \\ iii)\ \forall S' \neq S \wedge A \cap B = \varnothing \wedge A \not\to B,\\ \quad \dfrac{|B|/|S'|}{|A|/|S|} = \dfrac{P'(b)}{P(a)} = \dfrac{1}{P(a)} \dfrac{P(b)}{1 - P(a_i)},\ a_i \in A \\ iv)\ \forall S' \neq S \wedge A \cap B \neq \varnothing \wedge A \to B',\\ \quad \dfrac{|B'|/|S'|}{|A|/|S|} = \dfrac{P''(b)}{P(a)} = \dfrac{1}{P(a)} \dfrac{P(b) - P(b_i)}{1 - P(b_i)},\ b_i \in B \end{cases}
\end{aligned} \tag{10}$$

■

Example 7. In the *invariant* sample space $S_1 = \{H = 0.68,\ T = 0.32\}$ as modeled in Example 2 in Part I, the events *head* (*H*) and *tail* (*T*) are mutually exclusive in a *single* toss of the unfair coin. Therefore, the following probability divisions of unrelated events can be obtained according to Theorem 4(i), respectively:

$$P(H/T) = 0$$
$$P(T/H) = 0$$

Example 8. Redo Example 5 with none-mutually-exclusive events in $S_1 = \{H = 0.68,\ T = 0.32\}$, the following probability divisions between those of two consecutive tosses and the first toss can be obtained according to Theorem 4(ii), respectively, as follows:

$$P(HT/H) = \frac{P(HT)}{P(H)} = \frac{0.22}{0.68} = 0.32$$
$$P(TH/T) = \frac{P(TH)}{P(T)} = \frac{0.22}{0.32} = 0.69$$
$$P(HH/H) = \frac{P(HH)}{P(H)} = \frac{0.46}{0.68} = 0.68$$
$$P(TT/T) = \frac{P(TT)}{P(T)} = \frac{0.10}{0.32} = 0.31$$

It is noteworthy that, according to Theorem 4(ii), the event of the divisor must not be mutually exclusive to that of the dividend. Otherwise, Theorem 4(i) should be applied such as in the cases of $P(HH/T) = 0, P(TT/H) = 0, P(HT/T) = 0,$ and $P(TH/H) = 0$ in the given context.

Example 9. Given a *variant* sample space $S_2^{2'} = \{BW = 0.28,\ WB = 0.27,\ BB = 0.27,\ WW = 0.18\}$ as modeled in Example 5 in Part I, i.e., $S_2^{2'} \neq S_2^2$, the following probability divisions between two draws of the uneven balls in the bag can be obtained according to Theorem 4(iii) or 4(iv), respectively:

$$P(BW\ /\ B) = \frac{P'(BW)}{P(B)} = \frac{0.28}{0.55} = 0.51$$
$$P(WB\ /\ W) = \frac{P'(WB)}{P(W)} = \frac{0.27}{0.45} = 0.60$$
$$P(BB\ /\ B) = \frac{P''(BB)}{P(B)} = \frac{0.27}{0.55} = 0.49$$
$$P(WW\ /\ W) = \frac{P''(WW)}{P(W)} = \frac{0.18}{0.45} = 0.40$$

The results obtained in Example 9 can be verified by applying the multiplication rules given in Eq. 9(iii) and 9(iv) as shown in the following example. This approach is particularly useful when the product probability is unknown.

Example 10. Redo Example 9 in $S_2^{2'} = \{BW = 0.28, WB = 0.27, BB = 0.27, WW = 0.18\}$ according to Eq. 9(iii) and 9(iv) obtaining the same results as follows:

$$P(BW\ /\ B) = P'(W) = \frac{P(W)}{1 - P(b_i)} = \frac{0.45}{1 - 0.11} = \frac{0.45}{0.89} = 0.51$$
$$P(WB\ /\ W) = P'(B) = \frac{P(B)}{1 - P(w_i)} = \frac{0.55}{1 - 0.09} = \frac{0.55}{0.91} = 0.60$$
$$P(BB\ /\ B) = P''(B) = \frac{P(B) - P(w_i)}{1 - P(b_i)} = \frac{0.55 - 0.11}{1 - 0.11} = \frac{0.44}{0.89} = 0.49$$
$$P(WW\ /\ W) = P''(W) = \frac{P(W) - P(w_i)}{1 - P(w_i)} = \frac{0.45 - 0.09}{1 - 0.09} = \frac{0.36}{0.91} = 0.40$$

In probability theory, it is often interested in predicating the odds of random outcomes about the ratio of the probabilities of an event's success and failure.

Definition 1. An *odd,* $\Theta(e)$, is a ratio between probabilities of an event e and its complement, or that of its success s_e and failure f_e, i.e.:

$$\forall e, s_e, f_e \in E \subseteq S,$$
$$\Theta(e) \triangleq \frac{P(e)}{1 - P(e)} = \frac{P(e)}{P(\bar{e})} = \frac{P(s_e)}{P(f_e)} \tag{11}$$

It is noteworthy that the value of odds is a nonnegative real number, i.e., $\Theta(e) \geq 0$, which may be great than 1.0 according to Definition 1.

2.5 The Addition Operator on Conjunctive Probabilities

Theorem 5. The *addition of probabilities* of two conjunctive events a or b in the sample space S in $\mathfrak{U}$, $P(a+b)$, is determined by the sum of the probabilities of $P(a)$ and $P(b)$ excluding that of the intersection $P(a \times b)$ given $a \in A \subset S$ and $b \in B \subset S'$, i.e.:

$$P(a+b) = P(A \vee B) = P(a) + P(b) - P(a)P(b\,|\,a)$$
$$= \begin{cases} i)\ \text{Invariant } S, \text{unrelated } \overline{R}, \text{and ME-dependent } D\text{:}\ S \times \overline{R} \times D \\ \quad P(a) + P(b) \\ ii)\ \text{Invariant } S, \text{related } R, \text{ and independent } \overline{D}\text{:}\ S \times R \times \overline{D} \\ \quad P(a) + P(b) - P(a)P(b) \\ iii)\ \text{Variant } S', \text{unrelated } \overline{R}, \text{and independent } \overline{D}\text{:}\ S' \times \overline{R} \times \overline{D} \\ \quad P(a) + P(b) - P(a)P'(b) \\ iv)\ \text{Variant } S', \text{related } R \text{ and dependent } D\text{:}\ S' \times R \times D \\ \quad P(a) + P(b) - P(a)P''(b) \end{cases} \tag{12}$$

Proof. Theorem 5 can be proven according to Definition 10 in Part I and Theorem 1 as follows:

$$\forall a, b, a \in A \subset S, \text{ and } b \in B \subset S',$$
$$P(a+b) = P(A \vee B) = \frac{|A \cup B|}{|S|}$$
$$= \frac{|A|}{|S|} + \frac{|B|}{|S|} - \frac{|A \cap B|}{|S|}$$
$$= P(a) + P(b) - P(a)P(b\,|\,a)$$
$$= \begin{cases} i)\ \ \forall S' = S \wedge A \cap B = \varnothing \wedge A \to (B = \varnothing, \text{ME}), \\ \quad P(a) + P(b) \\ ii)\ \ \forall S' = S \wedge A \cap B \neq \varnothing \wedge A \not\to B, \\ \quad P(a) + P(b) - P(a)P(b) \\ iii)\ \forall S' \neq S \wedge A \cap B = \varnothing \wedge A \not\to B, \\ \quad P(a) + P(b) - \frac{P(a)P(b)}{1 - P(a_i)},\ P(a_i) = \frac{P(a)}{|A|} \\ \quad = P(a) + P(b) - P(a)P'(b) \\ iv)\ \forall S \neq S \wedge A \cap B \neq \varnothing \wedge A \to B', \\ \quad P(a) + P(b) - P(a)\frac{P(b) - P(b_i)}{1 - P(b_i)},\ P(b_i) = \frac{P(b)}{|B|} \\ \quad = P(a) + P(b) - P(a)P''(b) \end{cases} \tag{13}$$

■

Example 11. Suppose a system encompasses two components C_1 and C_2 with estimated failure rates as $F_1 = 0.7$ and $F_2 = 0.3$, respectively, in an invariant sample space. The conjunctive probabilities for a system failure of either C_1 or C_2 can be determined according to Theorem 5(ii) as follows:

$$P(F_1 + F_2) = P(F_1) + P(F_2) - P(F_1)P(F_2)$$
$$= 0.7 + 0.3 - 0.7 \bullet 0.3$$
$$= 1.0 - 0.21 = 0.79$$

Example 12. Reuse the individual probabilities obtained in Example 2 in Part I in the invariant sample space $S_1 = \{H = 0.68, T = 0.32\}$. The following additions of conjunctive probabilities for expecting some mixed head and tail of an unfair coin in two tosses can be derived according to Theorem 5(i):

$$P(HT+TH) = P(HT) + P(TH) - P(HT)P(TH) = 0.24 + 0.24 - 0 = 0.48$$
$$P(HH+TT) = P(HH) + P(TT) - P(HH)P(TT) = 0.36 + 0.16 - 0 = 0.52$$

Example 13. Consider the *variant* sample space $S_2^{2'} = \{BW = 0.28, WB = 0.27, BB = 0.27, WW = 0.18\}$ as modeled in Example 5 in Part I where no ball will be returned into the bag after a draw. The following probability additions between two conjunctive draws of the uneven balls in the bag can be obtained according to Theorem 5(iii) or 5(iv), respectively:

$$P(B+W) = P(B) + P(W) - \frac{P(B)P(W)}{1-P(w_i)} = 0.55 + 0.45 - \frac{0.55 \bullet 0.45}{1-0.11} = 1 - \frac{0.25}{0.89} = 0.72$$
$$P(W+B) = P(W) + P(B) - \frac{P(W)P(B)}{1-P(w_i)} = 1 - \frac{0.45 \bullet 0.55}{1-0.09} = 1 - \frac{0.25}{0.91} = 0.73$$
$$P(B+B) = P(B) + P(B) - \frac{P(B)(P(B)-P(b_i))}{1-P(b_i)} = 1.1 - \frac{0.55(0.55-0.11)}{1-0.11} = 1.1 - \frac{0.24}{0.89} = 0.83$$
$$P(W+W) = P(W) + P(W) - \frac{P(W)(P(W)-P(w_i))}{1-P(w_i)} = 0.9 - \frac{0.45(0.45-0.09)}{1-0.09} = 0.9 - \frac{0.16}{0.91} = 0.72$$

2.6 The Subtraction Operator on Decompositive Probabilities

The algebraic operation of probability subtraction is an inverse operation of probability addition, which is not defined in traditional probability theory.

Theorem 6. The *subtraction of related probability* of an event *b* from that of *a* in the sample spaces *S* in $\mathfrak{U}$, $P(a-b)$, is determined by the probability of event *a* excluding that of *b* given $a \in A \subset S$ and $b \in B \subset S'$, i.e.:

$$P(a-b) \triangleq P(a) - P(a)P(b \mid a) = \begin{cases} i) \text{ Invariant } S, \text{unrelated } \overline{R}, \text{ and ME dependent } D\text{: } S \times \overline{R} \times D \\ \quad \tilde{P}(\tilde{a}) \\ ii) \text{ Invariant } S, \text{ related } R, \text{ and independent } \overline{D}\text{: } S \times R \times \overline{D} \\ \quad P(a) - P(a)P(b) = P(a)P(\overline{b}) \\ iii) \text{ Variant } S', \text{ unrelated } \overline{R}, \text{ and independent } \overline{D}\text{: } S' \times \overline{R} \times \overline{D} \\ \quad P(a) - P(a)P'(b) = P(a)P'(\overline{b}) \\ iv) \text{ Variant } S', \text{ related } R \text{ and dependent } D\text{: } S' \times R \times D \\ \quad P(a) - P(a)P''(b) = P(a)P''(\overline{b}) \end{cases}$$

where $P(\overline{b}) = 1 - P(b)$. (14)

Proof. Theorem 6 can be proven according to Definition 10 in Part I and Theorem 1 as follows:

$$\forall a, b, a \in A \subset S, \text{ and } b \in B \subset S',$$
$$P(a-b) = P(A \setminus B) = \frac{|A \setminus B|}{|S|} = \frac{|A|}{|S|} - \frac{|A \cap B|}{|S|} = P(a) - P(b \mid a)$$
$$= \begin{cases} i) \ \forall S' = S \wedge A \cap B = \varnothing \wedge A \rightarrow (B = \varnothing, \text{ME}), \\ \quad P(a) \\ ii) \ \forall S' = S \wedge A \cap B \neq \varnothing \wedge A \not\rightarrow B, \\ \quad P(a) - P(a)P(b) = P(a)(1-P(b)) = P(a)P(\overline{b}) \\ iii) \ \forall S' \neq S \wedge A \cap B = \varnothing \wedge A \not\rightarrow B, \\ \quad P(a) - \frac{P(a)P(b)}{1-P(a_i)}, \ a_i \in A \\ \quad = P(a)(1 - \frac{P(a)P(b)}{1-P(a_i)}) = P(a)(1-P'(b)) \\ \quad = P(a)P'(\overline{b}) \\ iv) \ \forall S' \neq S \wedge A \cap B \neq \varnothing \wedge A \rightarrow B, \\ \quad P(a) - P(a)\frac{P(b)-P(b_i)}{1-P(b_i)}, \ b_i \in B \\ \quad = P(a)(1 - \frac{P(b)-P(b_i)}{1-P(b_i)}) = P(a)(1-P''(\tilde{b})) \\ \quad = P(a)P''(\overline{b}) \end{cases} \quad (15)$$

■

Example 14. Given the invariant sample space $S_1 = \{H = 0.68, T = 0.32\}$ as modeled in Example 2 in Part I, the following probability subtraction operations on the unfair coin can be derived according to Theorem 6(i) and 6(ii), respectively:

$$P(H-T) = P(H) - P(HT) = 0.68 - 0 = 0.68 \quad // ME$$
$$P(T-H) = P(T) - P(TH) = 0.32 - 0 = 0.32 \quad // ME$$
$$P(S_1 - H) = P(S_1)P(\overline{H}) = 1 \bullet (1-0.68) = 0.32 \quad // H \subset S_1$$
$$P(S_1 - T) = P(S_1)P(\overline{T}) = 1 \bullet (1-0.32) = 0.68 \quad // T \subset S_1$$
$$P(H-H) = P(\varnothing) = 0$$
$$P(T-T) = P(\varnothing) = 0$$

Example 15. Consider the *variant* sample spaces $S_2 = \{\mathop{R}\limits_{i=1}^{5} P(b_i \mid b_i \in B) = 0.11, \ \mathop{R}\limits_{i=6}^{10} P(w_i \mid w_i \in W) = 0.09\}$ and $S_2^{2'} = \{BW = 0.28, WB = 0.27, BB = 0.27, WW = 0.18\}$, respectively, as modeled in Examples 3 and 5 in Part I. The following probability subtraction operations on the uneven balls in the bag can be solved according to Theorem 6(iii), respectively:

$$P(B-W)=P(B)P'(\overline{W})=P(B)(1-\frac{P(W)}{1-P(b_i)})$$
$$=0.55(1-\frac{0.45}{1-0.11})=0.55\bullet 0.49=0.27$$
$$P(W-B)=P(W)P'(\overline{B})=P(W)(1-\frac{P(B)}{1-P(w_i)})$$
$$=0.45(1-\frac{0.55}{1-0.09})=0.45\bullet 0.40=0.18$$
$$P(BB-W)=P(BB)P''(\overline{W})=P(BB)(1-\frac{P(W)}{1-P(b_i)})$$
$$=0.27(1-\frac{0.45}{1-0.11})=0.27\bullet 0.49=0.13$$
$$P(WW-B)=P(WW)P''(\overline{B})=P(WW)(1-\frac{P(B)}{1-P(w_i)})$$
$$=0.18(1-\frac{0.55}{1-0.09})=0.18\bullet 0.40=0.07$$

Example 16. Given the same layout as that of Example 15, the following probability subtraction operations on the uneven balls in the bag can be solved according to Theorem 6(iv), respectively:

$$P(BW-B)=P(BW)P''(\overline{B})=P(BW)(1-\frac{P(B)-P(b_i)}{1-P(b_i)})$$
$$=0.28(1-\frac{0.55-0.11}{1-0.11})=0.28\bullet 0.51=0.14$$
$$P(WB-W)=P(WB)P''(\overline{W})=P(WB)(1-\frac{P(W)-P(w_i)}{1-P(w_i)})$$
$$=0.27(1-\frac{0.45-0.09}{1-0.09})=0.27\bullet 0.60=0.16$$
$$P(BB-B)=P(BB)P''(\overline{B})=P(BB)(1-\frac{P(B)-P(b_i)}{1-P(b_i)})$$
$$=0.27(1-\frac{0.55-0.11}{1-0.11})=0.27\bullet 0.51=0.14$$
$$P(WW-W)=P(WW)P'(\overline{W})=P(WW)(1-\frac{P(W)-P(w_i)}{1-P(w_i)})$$
$$=0.18(1-\frac{0.45-0.09}{1-0.09}=0.18\bullet 0.60=0.11$$

Table 2. Algebraic Rules of Probability Algebra

No.	Rule	Invariant sample space ($S'=S$)		Variant sample space
		Unrelated events ($A\cap B=\varnothing$)	**Related events** ($A\cap B\neq\varnothing$)	($S'\neq S$)
1	Commutative	$P(b\mid a)\neq P(a\mid b)$	$\neq$	$\neq$
		$P(a\times b)=P(b\times a)$	$=$	
		$P(a/b)\neq P(b/a)$	$\neq$	
		$P(a+b)=P(b+a)$	$=$	
		$P(a-b)\neq P(b-a)$	$\neq$	
2	Associative	$P(a\mid(b\mid c))\neq P((a\mid b)\mid c)$	$\neq$	$\neq$
		$P(a\times(b\times c))=P((a\times b)\times c)$	$=$	
		$P(a/(b/c))\neq P((a/b)/c)$	$\neq$	
		$P(a+(b+c))=P((a+b)+c)$	$=$	
		$P(a-(b-c))\neq P((a-b)-c)$	$\neq$	
3	Distributive	$P(a\times(b+c))=P((a\times b)+(a\times c))$	$=$	$\neq$
		$P(a\times(b-c))=P((a\times b)-(a\times c))$	$=$	
		$P((b+c)/a)=P((b/a)+(c/a)),\ P(a)>0$	$=$	
		$P((b-c)/a)=P((b/a)-(c/a)),\ P(a)>0$	$=$	
4	Transitive	$P(a)=P(b)\wedge P(b)=P(c)\Rightarrow P(a)=P(c)$	$=$	$=$
5	Complement	$P(\overline{a})=1-P(\overline{a}),\quad P(\overline{a})+P(\overline{a})=1$ $P(S)=1,\quad P(\varnothing)=0$ $P(\overline{S})=0,\quad P(\overline{\varnothing})=1$		
6	Involution	$P(\overline{\overline{a}})=P(a)$		
7	Idempotent	$P(a\times a)=P(a),\quad P(a+a)=P(a)$ $P(a/a)=1,\quad P(a-a)=0$		
8	Identity	$P(a\times S)=P(a),\quad P(a\times\varnothing)=0$ $P(a/S)=P(a),\quad P(S/a)=P^{-1}(a)$ $P(a+S)=1,\quad P(a+\varnothing)=P(a)$ $P(S-a)=P(S)-P(a),\quad P(\varnothing/a)=0$ $P(a-\varnothing)=P(a),\quad P(\varnothing-a)=0$		

Corollary 3. The *complement of probability* on an event $a \in A \subset S$ in $\mathfrak{U}$, $P(\overline{a})$, is a special case of probability subtraction, i.e.:

$$\begin{aligned} P(\overline{a}) &= 1 - P(\overline{a}) \\ &= P(S) - P(a) = P(S-a),\ a \in E \subseteq S \end{aligned} \quad (16)$$

3. The Formal Properties and Rules of the Extended Probability Theory

The mathematical model of general probability, the framework of the revisited probability theory, and the formal operators of probability algebra as developed in Sections 2 and Part I of the paper enable rigorous analyses of the nature, properties, and rules of probabilities as well as their algebraic operations. A set of 36 algebraic properties and rules of the extended probability operations is summarized in Table 2

Basic rules of probability algebra in the universe of discourse of probability $\mathfrak{U}$ can be expressed in categories of the *commutative, associative, distributive, transitive, complement, involution, idempotent,* and *identity* rules. It is noteworthy that it is unnecessary that each of the probability operators obeys all the general algebraic rules. Each algebraic rule on probability multiplication, division, addition, subtraction, conditional, and complement operations can be proven by applying related definitions and arithmetic principles. The algebraic rules of the probability theory may be applied to derive and simply complex probability operations in formal probability manipulations and uncertainty reasoning by both humans and cognitive systems. The framework of the revisited probability theory reveals that classic probability theory is a special case and subsystem of the revisited probability theory in terms of both mathematical models and probability operations.

4. Conclusion

As the second part of the revisited probability theory, a general theory of probability has been rigorously introduced as an extension of the classic probability to deal with complicated variant sample spaces as well as complex event relations and dependencies. The revisited probability theory has been formally described as a framework of hyperstructures of dynamic probability and their algebraic operations. Mathematical models and formal operators of the general probability framework have enabled rigorous analyses of the nature, properties, and rules of probability theories and their algebraic operations. It has been found that the conditional probability played a centric role in the framework of probability theories in order to solve the highly coupled cyclic-definition problems in traditional probability theories. It has been proven that Bayes' law may be revisited and validated based on the properties of the variant sample spaces as revealed in this paper. This work has also led to a theory of fuzzy probability that further extends the general probability theory to fuzzy probability spaces and fuzzy algebraic operations.

Acknowledgment: The author would like to acknowledge the support in part by a discovery fund of the Natural Sciences and Engineering Research Council of Canada (NSERC). The author would like to thank the anonymous reviewers for their valuable suggestions and comments on the previous version of this paper.

References:

[1] Artin, M. (1991), *Algebra*, Prentice-Hall, Inc., Englewood Cliffs, USA.

[2] Bender, E.A. (1996), *Mathematical Methods in Artificial Intelligence*, Los Alamitos, CA: IEEE CS Press.

[3] BISC (2013), *Internal Communications*, University of California, Berkeley, CA.

[4] Cantor, G. (1874), On a Characteristic Property of all Real Algebraic Numbers, *J. Reine Angew. Math.,* 77, 258-262.

[5] di Finetti, B. (1970), *Theory of Probability*, John Wiley & Sons, Inc., NY.

[6] Gowers, T. ed. (2008), *The Princeton Companion to Mathematics*, Princeton University Press, NJ.

[7] Hacking, I. (1975), *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Inductions and Statistical Inference*, Cambridge University Press, Cambridge, UK.

[8] Johnson, R.A. and G.K. Bhattacharyya (1996), *Statistics: Principles and Methods*, 3rd ed., John Wiley & Sons, Inc., NY.

[9] Kolmogorov, A.N. (1933), *Foundations of the Theory of Probability*, 2nd English ed. (1956), Chelsea Publishing Co., NY.

[10] Lipschutz, S. and M. Lipson (1997), *Theory and Problems of Discrete Mathematics*, 2nd ed., McGraw-Hill, NY.

[11] Mosteller, F. (1987), *Fifty Challenging Problems in Probability*, Dover Publishing, Inc., NY.

[12] Novak, V., I. Perfilieva, and J. Mockor (1999), Mathematical Principles of Logic, Kluwer, Boston/Dordrecht.

[13] Pedrycz, W. and F. Gomide (1998), Introduction to Sets, Cambridge, MA: MIT Press.

[14] Potter, M. (2004), *Set Theory and Its Philosophy: A Critical Introduction,* Oxford University Press.
[15] Ross, T.J. (1995), *Logic with Engineering Applications*, McGraw-Hill Co., 1995.
[16] Todhunter, I. (1865), *A History of the Theory of Probability from the Time of Pascal to that of Laplace*, Reprinted 1965, Chelsea Publishing Co., NY.
[17] Venn, J. (1888), *The Logic of Chance*, 3rd ed., Reprinted 1962, Chelsea Publishing Co., NY.
[18] Wang, Y. (2007), *Software Engineering Foundations: A Software Science Perspective*, CRC Series in Software Engineering, Vol. II, Auerbach Publications, NY, USA.
[19] Wang, Y. (2008a), On Contemporary Denotational Mathematics for Computational Intelligence, *Trans. Computational Science,* 2, 6-29.
[20] Wang, Y. (2008b), On the Big-R Notation for Describing Iterative and Recursive Behaviors, *International Journal of Cognitive Informatics and Natural Intelligence,* 2(1), Jan., 17-28.
[21] Wang, Y. (2011), Inference Algebra (IA): A Denotational Mathematics for Cognitive Computing and Machine Reasoning (I), *International Journal of Cognitive Informatics and Natural Intelligence,* 5(4), 61-82.
[22] Wang, Y. (2012a), In Search of Denotational Mathematics: Novel Mathematical Means for Contemporary Intelligence, Brain, and Knowledge Sciences, *Journal of Advanced Mathematics and Applications,* 1(1), 4-25.
[23] Wang, Y. (2012b), Formal Rules for Causal Analyses and Inferences, *International Journal on Software Science and Computational Intelligence*, 4(4) , 70-86.
[24] Wang, Y. (2012c), Inference Algebra (IA): A Denotational Mathematics for Cognitive Computing and Machine Reasoning (II), *International Journal of Cognitive Informatics and Natural Intelligence,* 6(1), 21-46.
[25] Wang, Y. (2014a), The Theory of Arithmetic in the Extended Domain of Numbers, *Journal of Advanced Mathematics and Applications,* 3(2), 165-175.
[26] Wang, Y. (2014b), Causal Inferences based on Semantics of Concepts in Cognitive Computing, *WSEAS Transactions on Computers*, 13, 430-441.
[27] Wang, Y. (2014c), Towards a Theory of Probability for Cognitive Computing, Proc. 13th IEEE International *Conference on Cognitive Informatics and Cognitive Computing* (ICCI*CC 2014), London, UK, IEEE CS Press, Aug., pp. 21-29.
[28] Wang, Y. (2014d), Keynote: Latest Advances in Neuroinformatics and Systems, *Proceedings of 2014 International Conference on Neural Networks and Systems,* Venice, Italy, March, pp. 14-15.
[29] Wang, Y. (2015a), Towards the Abstract System Theory of System Science for Cognitive and Intelligent Systems, *Springer Journal of Complex and Intelligent Systems,* June, 1(3), 11-32.
[30] Wang, Y. (2015b), Formal Cognitive Models of Data, Information, Knowledge, and Intelligence, *WSEAS Transactions on Computers*, 14, 770-781.
[31] Wang, Y. (2015c), Keynote: Cognitive Robotics and Mathematical Engineering, *14th IEEE International Conference on Cognitive Informatics and Cognitive Computing* (ICCI*CC 2015), Tsinghua University, Beijing, China, IEEE CS Press, July, pp. 4-6.
[32] Wang, Y. (2015d), Towards a Logical Algebra (FLA) for Formal Inferences in Cognitive Computing and Cognitive Robotics, *14th IEEE International Conference on Cognitive Informatics and Cognitive Computing* (ICCI*CC 2015), Tsinghua University, Beijing, China, IEEE CS Press, July, pp. 24-32.
[33] Wang, Y. (2015e), Fuzzy Probability Algebra (FPA): A Theory of Fuzzy Probability for Fuzzy Inference and Computational Intelligence, *Journal of Advanced Mathematics and Applications*, 4(1), 38-55.
[34] Whitworth, W.A. (1959), *Choice and Chance*, Hafner Publishing Co.
[35] Zadeh, L.A. (1965), Sets, *Information and Control*, 8, 338-353.
[36] Zadeh, L.A. (1968), Probability Measures of Events, *Journal of Math. Analysis and Appl.* 23, 421-427.
[37] Zadeh, L.A. (1996), Logic and the Calculi of Rules and Graphs, *Multiple-Valued Logic, 1*, 1-38.
[38] Zadeh, L.A. (2002), Toward a Perception-based Theory of Probabilistic Reasoning with Imprecise Probabilities**,** *Journal of Statistical Planning and Inference*, Elsevier Science, 105, 233-264.