# Recognition and Resolution of "Comprehension Uncertainty" in *AI*

Sukanto Bhattacharya[1,*] and Kuldeep Kumar[2]
*[1]Deakin Graduate School of Business, Deakin University,*
*[2]School of Business, Bond University,*
*Australia*

## 1. Introduction

### 1.1 Uncertainty resolution as an integral characteristic of intelligent systems

Handling uncertainty is an important component of most intelligent behaviour – so uncertainty resolution is a key step in the design of an artificially intelligent decision system (Clark, 1990). Like other aspects of intelligent systems design, the aspect of uncertainty resolution is also typically sought to be handled by emulating natural intelligence (Halpern, 2003; Ball and Christensen, 2009). In this regard, a number of computational uncertainty resolution approaches have been proposed and tested by Artificial Intelligence (*AI*) researchers over the past several decades since birth of *AI* as a scientific discipline in early 1950s post- publication of Alan Turing's landmark paper (Turing, 1950).

The following chart categorizes various forms of uncertainty whose resolution ought to be a pertinent consideration in the design an artificial decision system that emulates natural intelligence:
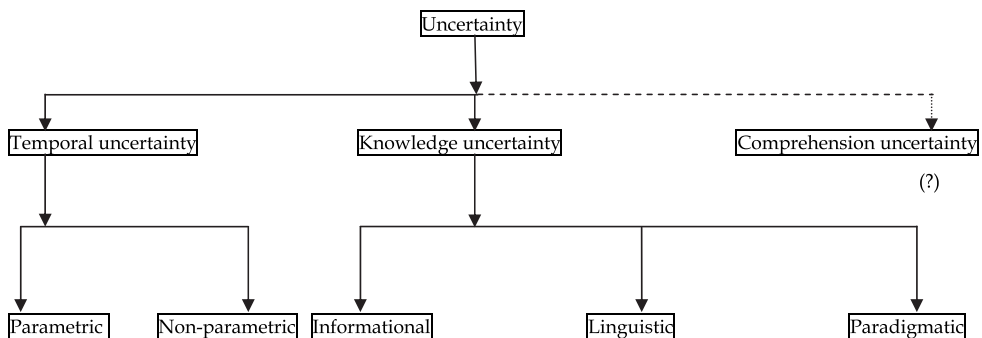


Fig. 1. Broad classifications of "uncertainty" that intelligent systems are expected to resolve

* Corresponding author

Temporal uncertainty, as the name suggests, arises out of *imperfect foresight* – i.e. it concerns the general problem of determining the future decision state of a dynamic system the current and past    decision states of which are known. As a sub-category of temporal uncertainty, parametric uncertainty is that form of uncertainty the resolution of which wholly depends on estimating a set of underlying parameters that determine a future decision state of a system given its current and/or past decision states. The fundamental premise is that there exist parameters, which if estimated accurately, would fully explain the temporal transition from current to a future decision state. In most practical *AI* applications it is handled by embedding an efficient parameter estimation *kernel* e.g. an asset price prediction kernel that is embedded within an intelligent financial trading system (Huang, Pasquier and Quek, 2009). On the other hand non-parametric uncertainty is that form of temporal uncertainty the resolution of which is either wholly or substantially independent of any parameters that can be statistically estimated from the current or past decision states of the system. That is, in resolving non-parametric uncertainty one cannot assume that there is a set of parameters whose accurate estimation can fully explain the dynamic system's time-path (Kosut, Lau and Boyd, 1992). To resolve non-parametric uncertainty, *AI* models are usually equipped with some feedback/learning mechanism coupled with a *performance measure index* that indicates when optimal learning has occurred so that predictive utility isn't lost on account of *overtraining* when predicting a future state using the current/past states as the *inputs* (Yang et al, 2010).

Knowledge uncertainty, again as the name suggests, arises out of *imperfect understanding* – i.e. it concerns the general problem of determining the future decision state of a dynamic system the knowledge about whose current and/or past states are either *incomplete*, *ill-defined* or *inconsistent*. If there is incomplete information available about the current decision state of the system then the sub-category of knowledge uncertainty it would be categorized under is informational uncertainty. A common way of dealing with informational uncertainty is to try and *enhance* the current level of information by applying an appropriate information theoretic tool e.g. Ding *et al* (2008) applied rough sets theory coupled with a self-adaptive algorithm to separately "mine" consistent and inconsistent decision rules; along with experimental validation for large incomplete information systems. If the information available about the current decision state of the system is ill-defined i.e. it is subject to *interpretational ambiguity* then it would come under the sub-category of linguistic uncertainty. A large part of interpretational ambiguity arises as a direct result of statements made in natural language (Walley and Cooman, 2001). Lotfi Zadeh, the proponent of *fuzzy logic*, contended that *possibility measures* are best used to resolve linguistic uncertainty in decision systems (Zadeh, 1965). If the information available about the current decision state of the system is inconsistent i.e. it is fundamentally dependent on the origin, then the resulting uncertainty would come under the sub-category of paradigmatic uncertainty. If available information is dependent on its origin then it can be expected to materially change if one chooses a different source for the same information. For example, software agents have to reason and act on a domain in which the universe of possible scenarios is fundamentally prescribed by the available metadata records. But these metadata records can sometimes be found to be mutually inconsistent when compared. The paradigmatic uncertainty resulting from the inconsistency and imprecision is best addressed by building in enough flexibility in the system so that the cogency of information related to the current

(and past) decision states gleaned from different sources is a *set-valued* rather than *point-valued* feature (Sicilia, 2006).

A three-valued extension of classical (i.e. binary) fuzzy logic was proposed by Smarandache (2002) when he coined the term "neutrosophic logic" as a generalization of fuzzy logic to such situations where it is impossible to *de-fuzzify* the original fuzzy–valued variables via some tractable membership function into either of set $T$ or its complement $T^C$ where both $T$ and $T^C$ are considered crisp sets. In these cases one has to allow for the possibility of a third unresolved state intermediate between $T$ and $T^C$. As an example one may cite the well known "thought experiment" in quantum metaphysics of *Schrödinger's cat* (Schrödinger, 1935) – the cat in a closed box is in limbo between two states "dead" and "alive" and it is impossible to tell which unless one opens the box at which point the effect of observer participation is said to intervene and cause that indeterminate state to collapse into a classical state of either a dead or an alive cat to be observed in the box. But as long as observer participation is completely absent one cannot in any way *disentangle* these two crisp sets!

This brings us to the final form of uncertainty that an artificially intelligent decision system ought to be able to resolve – something which we christened here as "comprehension uncertainty". While some elements of "comprehension uncertainty" is sought to be handled (often unknowingly) by the designers of intelligent systems by using one or more tools targeted to resolve either temporal or knowledge uncertainty, the concept of "comprehension uncertainty" has not yet been adequately described and addressed in contemporary *AI* literature. That is the reason we decided to depict this form of uncertainty using a *dashed rather than continuous connector* in the above chart. Also the question mark in the chart denotes the fact that there is no known repository of theoretical knowledge (not necessarily limited to the discipline of *AI*) that addresses such a form of uncertainty. The purpose of this chapter is to therefore posit a scientific theory of "comprehension uncertainty".

## 2. The meaning of "comprehension uncertainty"

While all the other forms of uncertainty as discussed above necessarily *originates from and deals with the contents/specification of an elementary set of interest*, which is a subset of the universal set, by the term "comprehension uncertainty" we mean and include any form of uncertainty that *originates from and deals with the contents/specification of the universal set itself*. If the stock of our entire knowledge about a problem is *universal* (i.e. there is absolutely nothing else that is 'fundamentally unknown' about that problem) only then we can claim to fully comprehend the problem so that no "comprehension uncertainty" would then exist. There is a need here to distinguish between "complete knowledge" and "universal knowledge". The knowledge about a problem can be said to be complete if it consists of the *entire stock of <u>current</u> knowledge* that is pertinent to that particular problem. However the current stock of knowledge, even in its entirety, may not be the universal knowledge simply because ways of adding to that current stock of knowledge could be beyond the current limits of comprehension i.e. *the universal set could itself be ill-defined.* If intelligent systems are primarily intended to emulate natural intelligence and treat "functional comparability" with natural intelligence as the most desirable outcome, then the limits to comprehension for natural intelligence should translate to similar limits for such systems as well.

## 2.1 How does natural intelligence resolve "comprehension uncertainty" in decision-making?

As highly evolved, intelligent beings, humans have become adept at continually taking decisions based on information that is subject to various forms of uncertainty. We can negotiate a busy sidewalk more often than not without colliding with other pedestrians and can cross a road safely (again most of the times) without being flattened by a car although we have at best a very imprecise idea of the speed of an oncoming car. Human brain, as the highest seat of natural intelligence, has evolved unique ways of working with various uncertainties including "comprehension uncertainty". Humans are also dealing with "comprehension uncertainty", for example when designing an unmanned, deep-space probe. We design the space probe using our current stock of knowledge in astrophysics; thermodynamics etc., identifying, assessing and resolving the pertinent temporal and knowledge uncertainties. At the same time we are also cognisant of a *gap* in our knowledge. This is not because we haven't been able to fully utilize our current stock of knowledge; rather it is the gap that exists between our current knowledge of deep space etc. and the *universal* knowledge which is outside of our "limits" of comprehension i.e. primarily originating from an ill-defined universal set.

Artificially intelligent decision systems are typically programmed to inexorably seek a 'global' optimum while in reality, the presence of "comprehension uncertainty" will always negate that prospect. What an intelligent system returns as a 'global' optimum is thus at best only such within its current domain knowledge and not a "universal" optimum. But an artificially intelligent system will always terminate its search once it attains what it perceives as the "global" optimum; based on the underlying premise that its current stock of domain-specific knowledge is in fact the universal one! On the other hand, naturally intelligent beings recognize the fundamental gap between current and universal knowledge and so will endeavour to keep expanding their "limits of comprehension".

An artificially intelligent decision system ought to be designed to 'realize' that its current stock of knowledge may not be the universal knowledge pertinent to a decision problem it is invoked to work out. Emulating natural intelligence, *AI models* should aim to be 'auto-cognisant' of any fundamental knowledge gaps and therefore be able to reconcile any deviations of the "global" from the "universal" optimum. A first step towards that is *effective operationalization* of the "comprehension uncertainty" concept. In the following section we posit and develop a formal conceptualization of the "comprehension uncertainty" concept. This basically involves an extension of classical probability theory to a realm of *higher-order probabilities* in a manner that is computationally tractable and fully reconcilable with the classical theory. Finally we posit and defend a logical framework justifying the due consideration of "comprehension uncertainty" in the context of designing artificially intelligent systems for practical applications in business, industry and society.

## 3. Developing some necessary theoretical groundwork

The primary objective of our work here is to simply posit the logically conceivable underpinnings of a probability theory extended to formalize comprehension uncertainty – we believe that our main purpose here is to merely open the proverbial Pandora's Box and thereby spawn a healthy stream of new research along both philosophical as well as

mathematical lines. In that desired direction, we firstly posit and prove a fundamental theorem necessary for such an extension to the theory of probability. Subsequently we show some computational 'tests' to illustrate the posited framework.

### 3.1 A foray into higher order probabilities

It is well known that much of modern theory of probability rests upon the three fundamental *Kolmogorov axioms* (Kolmogorov, 1956) which are conventionally stated as follows:

1st axiom: The probability of any event is a non-negative real number i.e. $P(E) \geq 0 \quad \forall E \in U$

2nd axiom: The probability of any one of the elementary events in the whole event space occurring is 1 i.e. $P(U)=1$

3rd axiom: Any countable sequence of pair-wise *non-overlapping* events $E_1$, $E_2$, ... $E_n$ satisfies the following relation: $P(E_1 \cup E_2 \cup ... \cup E_n) = \sum_i P(E_i)$; i = 1, 2, ..., n.

It is basically Kolmogorov's second and third axioms as noted above that render any extensions of the probability concept to higher orders (i.e. "probability of probability") superfluous as the information content of any such higher order probability can be satisfactorily transmuted via existing set-theoretic constructs. So, extending to a higher order would arguably yield trivial information. However the Kolmogorov axioms by themselves are also open to 'extensions' – for instance there is previous research that has revisited the proofs of the well-known Bell inequality based on underlying assumptions of separability and non-contextuality and constructed a model of generalized "non-contextual contrapositive conditional probabilities" consistent with the results of the famous Aspect experiment showing in general such probabilities are not necessarily all positive (Atkinson, 2000). By themselves the Kolmogorov axioms do not unequivocally rule out an extension of the definition of the universal set U itself so as to make U possess a *time-dynamic* rather than a *time-static* nature. So; in effect this means that if we were to consider a time-dynamic version of the universal set; then one would suddenly find that the information content of higher order probability no longer remains trivial i.e. an extension of the probability concept to higher orders (i.e. "probability of probability") is no longer superfluous – in fact it is logical! The good thing is that no new probability calculus needs to be formulated to describe such a theory of higher-order probabilities and this extended theory could still rest on the Kolmogorov axioms and could still draw fundamentally from the standard set-theoretic approach (as we will be demonstrating shortly); by merely using an extended definition of the universal set U which would now denote not merely an event space but a broader concept, which we christen as *event-spacetime*, i.e. an event space that can *evolve* over a time dimension.

Perhaps the only academic work preceding ours to have alluded that a higher-order probability theory is justifiable by an event space evolving over time was that by Haddawy and others (Haddawy, 1996; Lehner, Laskey and Dubois, 1996), where they provided "a logic that incorporates and integrates the concepts of subjective probability, objective probability, time and causality" (Lehner, Laskey and Dubois, 1996). We take a similar philosophical stance but go on to explicitly develop a logically tenable higher-order probability concept in discrete time. We have no doubt that an extension in continuous time is also attainable but it's left for later.

**Lemma 1**

*The probability that any one of the elementary events contained within the event space-time will occur between two successive time points $t_0$ and $t_1$ given that the contents/contours of the event space remains unchanged from $t_0$ to $t_1$ is unity i.e. $P(U_0 | U_0 = U_1) = 1$. By extension, $P(U_t | U_t = U_{t+1}) = 1$ for all $t = 0, 1, 2, 3, ...$*

**Proof**

Lemma 1 results from a natural extension of Kolmogorov's second axiom if we allow the event space to be of a time-dynamic nature i.e. if U is allowed to evolve through time in discrete intervals.

*QED*

**Lemma 2**

*If the classical probability of occurrence of a specific elementary event E contained within the event space-time is defined as $P(E)$, then the first-order probability of occurrence of such event E becomes $P\{P(E)\} = P^1(E) = P\{E | (U_0 = U_1)\} = P(E).[P\{(U_0 = U_1) | E\}/P(U_0 = U_1)]$*

**Proof**

Applying the fundamental law of conditional probability we can write as follows:

$$P\{E | (U_0 = U_1)\} = P\{E \cap (U_0 = U_1)\}/ P(U_0 = U_1)$$

$P\{E \cap (U_0 = U_1)\} = P\{(U_0 = U_1) \cap E\} = P(E).P\{(U_0 = U_1) | E\}$; and thus the result follows.

*QED*

**Lemma 3**

*Given the first-order probability of occurrence of elementary event E and assuming that $(U_t = U_{t+1})$ and $(U_{t+1} = U_{t+2})$ are independent for all $t = 0, 1, 2, 3, ...$, the second-order probability of occurrence of E becomes $P^2(E) = P(E) . P^1(E). [P\{(U_1 = U_2) | E\}/P(U_1 = U_2)]$.*

**Proof**

By definition, $P^2(E) = P\{P^1(E)\} = P[\{E | (U_0 = U_1)\} \cap \{E | (U_1 = U_2)\}]$

Since $(U_t = U_{t+1})$ and $(U_{t+1} = U_{t+2})$ are assumed independent for $t = 0, 1, 2, 3, ...$, we can write:

$$P[\{E | (U_0 = U_1)\} \cap \{E | (U_1 = U_2)\}] = P\{E | (U_0 = U_1)\} . P\{E | (U_1 = U_2)\}.$$

Substituting $P\{E | (U_0 = U_1)\}$ with $P^1(E)$ and then applying the fundamental law of conditional probability; the result follows.

*QED*

Thus, given the first-order probability of occurrence of an elementary event E, the second-order probability is obtained as a "probability of the first-order probability" and is necessarily *either equal to or less than* the first-order probability, as is suggested by common intuition. This logic could then be extended to each of the subsequent higher order probability terms. Based on lemmas 1 – 3, we next propose and prove a fundamental theorem of higher order (hereafter *H-O*) probabilities.

**A fundamental theorem of higher order probabilities (in discrete time)**

*If we set $P^0(E) \equiv P(E)$, then $P^t(E) = P(E) . P^{t-1}(E).[P\{(U_{t-1} = U_t) \mid E\}/P(U_{t-1} = U_t)]$ for t = 1, 2,3, ..., n*

**Proof**

$$P^1(E) = P\{P^0(E)\} = P(E) . [P\{(U_0 = U_1) \mid E\}/P(U_0 = U_1)] \qquad \text{from lemma 2}$$

$$P^2(E) = P\{P^1(E)\} = P(E) . P^1(E) . [P\{(U_1 = U_2) \mid E\}/P(U_1 = U_2)] \quad .. \text{from lemma 3}$$

$$P^3(E) = P[\{E \mid (U_0 = U_1)\} \cap \{E \mid (U_1 = U_2)\} \cap \{E \mid (U_2 = U_3)\}]$$

$$= P[[\{E \mid (U_0 = U_1)\} \cap \{E \mid (U_1 = U_2)\}] \cap \{E \mid (U_2 = U_3)\}]$$

$$= P^2(E) . P\{E \mid (U_2 = U_3)\}$$

$$= P(E). P^2(E).[P\{(U_2 = U_3) \mid E\}/P(U_2 = U_3)$$

$$P^4(E) = P[\{E \mid (U_0 = U_1)\} \cap \{E \mid (U_1 = U_2)\} \cap \{E \mid (U_2 = U_3)\} \cap \{E \mid (U_3 = U_4)\}]$$

$$= P[[\{E \mid (U_0 = U_1)\} \cap \{E \mid (U_1 = U_2)\} \cap \{E \mid (U_2 = U_3)\}] \cap \{E \mid (U_3 = U_4)\}]$$

$$= P^3(E) . P\{E \mid (U_3 = U_4)\}$$

$$= P(E). P^3(E).[P\{(U_3 = U_4) \mid E\}/P(U_3 = U_4)]$$

Extending to the (t-1)-th term, we can therefore write:

$$P^{t-1}(E) = P(E) . P^{t-2}(E).[P\{(U_{t-2} = U_{t-1}) \mid E\}/P(U_{t-2} = U_{t-1}) \qquad (1)$$

The expression for the t-th term is derived from (1) as follows:

$$P^t(E) = P(E). P^{(t-2)+1}(E).[P\{(U_{(t-2)+1} = U_t) \mid E\}/P(U_{(t-2)+1} = U_t)$$

$$= P(E).P^{t-1}(E).[P\{(U_{t-1} = U_t) \mid E\}/P(U_{t-1} = U_t) \qquad (2)$$

However we may also write:

$$P^t(E) = P[\{E \mid (U_0 = U_1)\} \cap \{E \mid (U_1 = U_2)\} \cap \{E \mid (U_2 = U_3)\} \cap \{E \mid (U_3 = U_4)\} \cap ... \cap \{E \mid (U_{t-1} = U_t)\}]$$

$$= P[[\{E \mid (U_0=U_1)\} \cap \{E \mid (U_1=U_2)\} \cap \{E \mid (U_2=U_3)\} \cap \{E \mid (U_3=U_4)\} \cap ... \cap \{E \mid (U_{t-2} = U_{t-1})\} \cap \{E \mid (U_{t-1} = U_t)\}]$$

$$= P^{t-1}(E) . P\{E \mid (U_{t-1} = U_t)\}$$

$$= P^{t-1}(E). P(E). [P\{(U_{t-1}=U_t) \mid E\}/P(U_{t-1}=U_t)] \qquad (3)$$

As (2) is identical to (3); by *principle of mathematical induction* the general case is proved for t= n.

*QED*

Obviously then, if $P(U_{t-1}=U_t) = P\{(U_{t-1}=U_t) \mid E\}$, for all t = 1, 2, 3, ..., n; we will end up with $P^n(E) = [P(E)]^n$ which makes this approach to *H-O* probability fully consistent with classical probability theory and in fact a very natural extension thereof if one sees the fundamentally time-dynamic characteristic of U.

## 3.2 Simple computational 'tests' to better illustrate the above-posited concept of *H-O* probability

To provide a simple illustration of how the *H-O* probabilities would pan out in discrete event-spacetime we have done a series of computations the results of which are graphically represented below. The graphs show the temporal evolution of the event-spacetime in discrete "time steps" and the resulting $P^t(E)$ values for t = 1, 2, ..., 5. We assume three temporal evolution forms – "*expanding event-spacetime*", "*contracting event-spacetime*" and "*oscillating event-spacetime*" and plot the $P^t(E)$ values for each of these three forms starting with a pervading assumption that $P(U_{t-1} = U_t) = 1$. This assumption simplifies a lot of the computations as $P^t(E)$ then depends totally on $P\{(U_{t-1} = U_t)/E\}$. When $P\{(U_{t-1} = U_t)/E\} = 1$, we see that $P^t(E)$ converges to $P(E)^t$ for all values of t. On the other hand, when $P\{(U_{t-1} = U_t)/E\} = 0$, $P^t(E)$ converges to zero for all values of t. So, holding $P(E) = 0.10$, in an "expanding event-spacetime", $P^1(E) = P(E) = 0.10$, $p^2(E) = 0.10^2 = 0.01$ and so on for $P\{(U_{t-1}=U_t)/E\} = 1$. For $P\{(U_{t-1}=U_t)/E\} = 0$, $P^1(E) = p^2(E) = P^3(E) = p^4(E) = P^5(E) = 0$, while $P^t(E)$ values are seen to oscillate for $P\{(U_{t-1}=U_t)/E\}$ values randomly oscillating about 0.50 – the degree of oscillation decreasing with increasing order of probability i.e. $P^1(E)$ oscillates more than $P^2(E)$, $P^2(E)$ more than $P^3(E)$ and so on.
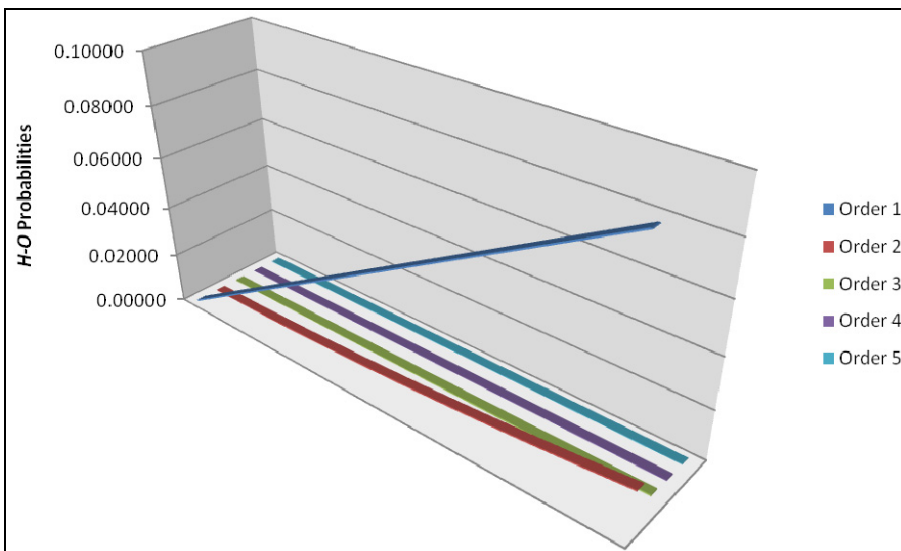


Fig. 2. Expanding Event-Spacetime, $[P(E) = 0.10]$
    Plot of $P^t(E)$; t = 1, 2, ..., 5 for $P\{(U_{t-1} = U_t)/E\}$ increasing from 0 to 1 in steps of 0.05

The expanding event-spacetime represents the situation where, with passage of time and evolution of the current stock of domain knowledge, there is a steadily increasing "probability of probability" of the occurrence of the elementary event of interest. The contracting event-spacetime represents the situation where, with passage of time and evolution of the current stock of domain knowledge, there is a steadily decreasing "probability of probability" of the occurrence of the elementary event of interest. The oscillating event-spacetime represents the situation where, with passage of time and

evolution of the current stock of domain knowledge, there is an erratic pattern in the "probability of probability" of the occurrence of the elementary event of interest because of the fact that some old knowledge that were 'replaced' by new knowledge make comebacks following newer discoveries.
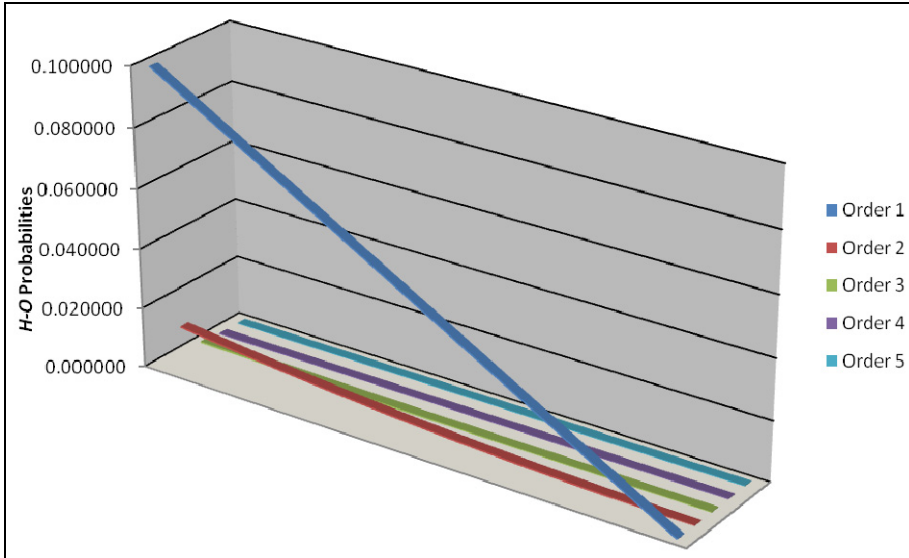


Fig. 3. Contracting Event-Spacetime, [P($E$) = 0.10]
       Plot of P$^t$($E$); t = 1, 2, ..., 5 for P{($U_{t-1}$ = $U_t$)/$E$} decreasing from 1 to 0 in steps of 0.05
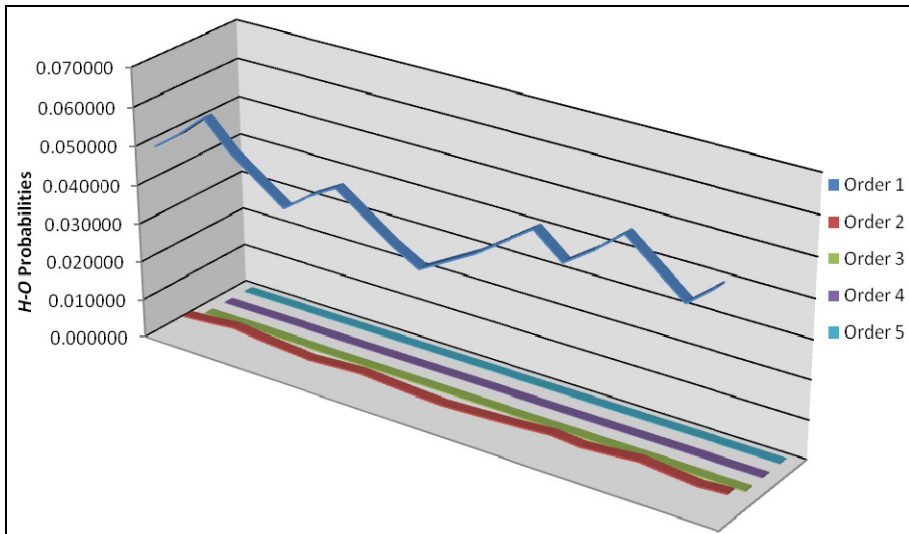


Fig. 4. Oscillating Event-Spacetime, [P($E$) = 0.10]
       Plot of P$^t$($E$); t = 1, 2, ..., 5 for P{($U_{t-1}$ = $U_t$)/$E$} allowed to randomly oscillate about 0.50

### 3.3 *H-O* probability implications for intelligent resolution of comprehension uncertainty

Although we do not mathematically compute *H-O* probabilities while taking decisions (or for that matter even ordinary probabilities), human intelligence does enough 'background processing' of fringe information (mostly even without knowing) to 'see' a bigger picture of the likely scenarios. Going back to the example of crossing a busy road, we are continuously processing information (often unknowingly) from the environment in terms of the rapidly changing pertinent event space. As long as the pertinent event space is 'pre-populated' with likely forms of road hazards, an artificially intelligent system can be 'trained' to emulate human decision-making and cross the road. It is when the contents of the pertinent event space dynamically changes that would throw off even the most advanced of AI-based systems given the current state of design of such systems. This is pretty much what Bhattacharya, Wang and Xu (2010) identified as a 'gap' in the current state of design of intelligent systems. The current design paradigm is overwhelmingly concerned with the "how" rather than the "why" – and resolution of comprehension uncertainty involves more of the "why". Rather than trying to answer "how to avoid being hit by a vehicle or some other hazard while crossing" *AI* designers ought to be focusing on "why are we vulnerable while crossing a busy road".

As soon as the focus of the design shifts to the "why", the link with comprehension uncertainty becomes a very natural extension thereof. Then we are simply asking <u>why</u> a particular event space is a pertinent one for the problem at hand? The natural answer is that in a specified time window, it contains all the elementary events out of which one or a few are conducive for the desired outcome. Then the question naturally progresses to what would happen outside that specified time window? If we are pre-populating the pertinent event space and then assuming that it would hold good for all times, it would be at the cost of ignoring comprehension uncertainty which can defeat the AI design. At this point it is perhaps useful to again remind readers that it is not the vagueness or imprecision associated with some contents of an event space that is of importance here (existing uncertainty resolution methods like rough sets, fuzzy logic etc. are adequate for dealing with those) – it is a temporal instability of the event space itself that is crux of the comprehension uncertainty concept.

The mathematics of *H-O* probabilities then offers a plausible route towards formal incorporation of comprehension uncertainty within artificially intelligent systems designed to replicate naturally intelligent decision-making. As naturally intelligent beings, humans are capable of somehow grasping the "limits to comprehension" that result from a gap between current knowledge and universal knowledge. If this was not the case then 'research' as an intellectual endeavour would have ceased! In the current design paradigm the focus is on training *AI* models to 'search' for global optimality while, ideally, the focus ought to be on training such models to do 'research' rather than 'search'! Recognition and incorporation of comprehension uncertainty in their learning framework would at least allow future *AI* models to 'grasp' the limits to comprehension so as not to invariably terminate as soon as a 'globally optimal' decision point has been reached using the current domain knowledge.

## 4. Conclusion: "comprehending the incomprehensible" – the future of *AI* systems design

In its current state, the design of artificially intelligent systems is pre-occupied with solving the "how" problems and as such do not quite recognize the need for resolving comprehension uncertainty. In fact, the concept of comprehension uncertainty was not even formally posited prior to this work by us although there have been a few takes on the mathematics of *H-O* probabilities. Earlier researchers mainly found the concept of *H-O* probabilities superfluous because they failed to view it in the context of formalizing comprehension uncertainty like we have done in this article.

However, given that the exact emulation of human intelligence continues to remain the Holy Grail for *AI* researchers, they have to grapple with comprehension uncertainty at some point or the other. The reason for this is simple – a hallmark of human intelligence is that it recognizes the limitations of the current stock of knowledge from which it draws. Thus any artificial system that ultimately seeks to emulate that intelligence must also necessarily see the limitations in current domain knowledge and allow for the fact that the current domain knowledge can evolve over time so that the global optimum attained with the current stock of knowledge may not remain the same at a future time. Once an artificially intelligent system is hardwired to recognize the *time-dynamic* aspect of the relevant event space within which it has to calculate the probabilities of certain outcomes and take a decision so as to maximize the expected value of the most desirable outcome, it will not terminate its search as soon as global optimality is reached in terms of the contents/contours of the current event space. It would rather go into a 'dormant' mode and continue to monitor the evolution of the event space and 're-engage' in its search as soon as $P\{(U_{t-1}=U_t)/E\} > 0$ at any subsequent time point.

With the formal hardwiring of comprehension uncertainty within the core design of an artificially intelligent system it can be trained to transcend from simply answering the "how" to ultimately formulating the "why" – firstly; why is the current body of knowledge an exhaustive source to draw from for finding the optimal solution to a particular problem and secondly; why that current body of knowledge may not be continue to remain an exhaustive source to draw from for all time in future. When it has been trained to formulate these "why" questions, only then can we expect an artificially intelligent system to take that significant leap towards finally gaining parity with natural intelligence.

## 5. References

Atkinson, D. (2000). Bell's Inequalities and Kolmogorov's Axioms, *Pramana – Journal of Physics*, Vol. 54, pp. 1–15

Ball, L. J. and B. T. Christensen (2009). Analogical reasoning and mental simulation in design: two strategies linked to uncertainty resolution, *Design Studies*, Vol. 30, No. 2, pp. 169–186

Bhattacharya, S., Y. Wang and D. Xu (2010). Beyond Simon's Means-Ends Analysis: Natural Creativity and the Unanswered 'Why' in the Design of Intelligent Systems for Decision-Making, *Minds and Machines*, Vol. 20, No. 3, pp.327–347

Clark, D. A. (1990). Numerical and symbolic approaches to uncertainty management in AI, *Artificial Intelligence Review*, Vol. 4, pp. 109–146

Ding, Z., X. Zhu, J. Zhao and H. Xu (2008). New knowledge acquisition method in incomplete information system based on rough set and self-adaptive genetic algorithm, In: *Proceedings of IEEE International Conference on Granular Computing*, Hangzhou, China (26th-28th August), pp. 196–200

Haddawy, P. (1996). Believing change and changing belief, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, Vol. 26, No. 3, pp.385–396

Halpern, J. Y. (2003). *Reasoning about uncertainty*, Cambridge, MA, and London: MIT Press

Huang, H., M. Pasquier and C. Quek (2009). Financial Market Trading System with a Hierarchical Co evolutionary Fuzzy Predictive Model, *IEEE Transactions on Evolutionary Computation*, Vol. 13, No. 1, pp. 56–70

Kolmogorov, A. N. (1956). *Foundations of the Theory of Probability*, Second Ed. (English), Chelsea Publishing Company, NY

Kosut, R. L., M. K. Lau and S. P. Boyd (1992). Set-Membership Identification of Systems with Parametric and Nonparametric Uncertainty, *IEEE Transactions on Automatic Control*, Vol. 37, No. 7, pp. 929–941

Lehner, P. E., K. B. Laskey and D. Dubois (1996). An introduction to issues in higher order uncertainty, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* Vol. 26, No. 3, pp.289–293

Schrödinger, E. (1935). Die gegenwärtige Situation in der Quantenmechanik (The present situation in quantum mechanics), *Naturwissenschaften*, Vol. 23, pp. 807–849

Smarandache, F. (2002). Preface: An Introduction to Neutrosophy, Neutrosophic Logic, Neutrosophic Set, and Neutrosophic Probability and Statistics, In: F. Smarandache (Ed.) *Proceedings of the First International Conference on Neutrosophy, Neutrosophic Logic, Neutrosophic Set, and Neutrosophic Probability and Statistics*, Xiquan, Phoenix, AZ

Sicilia, M-A. (2006). On Some Problems of Decision-Making under Uncertainty in the Semantic Web, *Studies in Fuzziness and Soft Computing*, Vol. 204, pp. 231–246

Turing, A. M. (1950). Computing machinery and intelligence, *Mind*, Vol. 59, pp. 433–460

Walley, P. and G. de Cooman (2001). A behavioral model for linguistic uncertainty, *Information Sciences*, Vol. 134, pp. 1–37

Yang, Z., S. C. P. Yam, L. K. Li and Y. Wang (2010). Universal Repetitive Learning Control for Nonparametric Uncertainty and Unknown State-Dependent Control Direction Matrix, *IEEE Transactions on Automatic Control*, Vol. 55, No. 7, pp. 1710 – 1715

Zadeh, L. A. (1965). Fuzzy sets, *Information and Control*, Vol. 8, pp. 338–353