The Z-number Enigma:
A Study through an Experiment

Romi Banerjee and Sankar K. Pal

Centre for Soft Computing Research
Indian Statistical Institute
Kolkata 700 108, India

(rm.banerjee@gmail.com, sankar@isical.ac.in)

Abstract
The Z-number, proposed by Zadeh in the year 2011, is a new fuzzy-theoretic approach to the Computing With Words (CWW) paradigm. It aspires to capture the uncertainty of information conveyed by a sentence, and serve as a model for the precisiation and linguistic summarization of a natural language statement. The Z-number thereby, lends a new dimension to CWW – uniting CWW with Natural Language Processing (NLP). This article is an illumination upon our exploration of the Z-number approach to CWW. Here, we enlist the probable contributions of the Z-number to CWW, present our algorithm for CWW using the Z-number, and describe a simulation of the technique with respect to a real-life example of CWW. In the course of the simulation, we extend the interpretation of the set-theoretic intersection operator to evaluate the intersection of perceptions and discover some of the challenges underlying the implementation of the Z-number in the area of CWW.

Keywords
Computing With Words (CWW), cognition, fuzzy sets, linguistics, machine learning, text summarization, natural language processing, perceptions, soft computing, natural computing

1. Introduction

Words encode perceptions and are inherently imprecise; not all events are expressible through precise numbers and symbols. Cognitive text or speech comprehension begins with the identification of the meaning of the constituent words, as per their usage, to arrive at the meaning of sentences; and a union of these sentence-perceptions leads to the comprehension of the complete text or speech sample.

The Computing With Words (CWW) paradigm, coined by Zadeh in 1996, draws inspiration from the remarkable perception-based decision-making ability of the human brain; the perceptions being encoded in the words and phrases used to describe events. The Intelligent Systems Revolution symbolizes the generation of machines that possess high levels of Machine IQ [1], [2], and undeniably, an implementation of the CWW paradigm would be a step in that direction.

Realization of the paradigm requires the machine to comprehend word-perceptions as well as a human being. Consequently, not only does the machine need to learn words – both existing and new, it also needs to apply them to form semantically, and ideally syntactically, correct natural language statements. Meanings of words, thus, need to be translated into some symbolic form. CWW is deeply engrained in the concepts of fuzzy logic [3], fuzzy linguistics [4], test score semantics [5] and PNL [6], and is the precursor to "Computing With Perceptions (CTP) [7]".

The Z-number [8], proposed by Zadeh, is a very new approach to the CWW paradigm. Besides unifying the concepts of fuzzy logic, fuzzy linguistics, test score semantics and PNL, the Z-number incorporates a measure of the reliability of the information in the sentence. Consequently, the Z-number could be sought as a medium of extension of the basic element of computation of CWW from word-perceptions to sentence-perceptions. We envision that this would lead to the development of a model of the natural process of comprehension in human beings.

In this article, we present a comprehensive study of the Z-number approach to CWW from an NLP perspective [9]. We begin with a recapitulation on sentences and its types (Section 2), followed by an overview of the basics of the Z-number (Section 3.1), and then move on to predict some of the contributions of the Z-number methodology to CWW (Section 3.2) and formulate a basic algorithm for the Z-number based CWW (Section 3.3).

The article, thereafter, focuses on our experiment where the Z-number approach is applied to a real-life example of CWW (an attempt at the simulation of the natural CWW in a machine). The experiment begins with an algorithm underlying the intuitive process of book selection in human beings (Section 4.1), which is then extended to include the Z-number components to simulate the book selection process in a machine (Section 4.2.B). The experiment calls for the extension of the basic intersection operator to allow the intersection of perceptions (Section 4.2.A) and leads to the discernment of some of the inherent challenges underlying the Z-numbers in CWW (Section 4.2.C).

2. A Recapitulation on Sentences [10]

A system that is to compute with words, essentially deals with computing with the words in sentences. In acknowledgement to the diversity of human speech and the different ways the same event can be described, this section prepares the base for the different kinds of sentences that such a system needs to anticipate and comprehend. Each type of sentence has its own unique set of rules and the component words and phrases need to be processed accordingly.

In the article, we use the words 'statement' and 'sentence' to mean the same – "Sentences are collections of words that make complete sense. The sense is not complete, unless something is being said about something".

Sentences are classified into a number of categories:

i. Based on the number of independent clauses –
   a. Simple sentences: Sentences with a single independent clause.
      E.g.: I love to read.
   b. Compound sentences: Sentences with two or more independent clauses.
      E.g.: He came back tired for he had been working all day.
   c. Complex sentences: Sentences with a single independent clause and more than one dependent clause.

E.g.: They went for a movie after they had completed their homework.

ii. Based on the nature of the sentences –

a. Declarative sentences: Sentences that are used to make a simple statement. A declarative sentence ends with a full stop.

E.g.: I love to read.

b. Interrogative sentences: Sentence that are used to make queries. An interrogative sentence ends with a question mark.

E.g.: Do you love to read?

c. Exclamatory sentences: Sentences that are used to emphasize a fact or convey an emotion. An exclamatory sentence ends with an exclamation mark.

E.g.: What a wonderful read!

d. Imperative sentences: Sentences that are used to command or request, with the pronoun 'you' implied.

E.g.: Please get me something to read.

e. Conditional sentences: Sentences that indicate dependencies between conditions, a "cause and effect" relationship.

E.g.: If I had a billion dollars, I would buy a castle of books.

Though not in common use, types of sentences also include –

i. Rhetorical Question: A question that is posed for persuasive effect without the expectation of an answer.

E.g.: When will people learn the consequences of turning every bit of open space into cement monstrosities?

ii. Paradox: A paradox is a logical statement or group of statements that lead to a contradiction or a situation which (if true) defies logic or reason.

E.g. : One thing that I know is that I know nothing (Socrates' Paradox).

3. The Z-number [8]

Actions rely on decisions which again depend on the information provided. Thus, greater the reliability of the information better is the decision made. The Z-number approach to CWW includes a measure of the reliability of the information in a statement, along with other parameters that result in the linguistic summarization of the sentence. The Z-number could thus form the basis of CWW-based discourse-oriented systems.

3.1 An Overview of the Z-number

Given a natural language statement, Y, the Z-number of Y is defined as a 2-tuple $Z = \langle A, B \rangle$, where A, a linguistic value, implies the restriction (constraint) on the values of X, a real valued linguistic variable – interpreted as the subject of Y, and B is a measure of the reliability (certainty) of A. Typically, A and B are expressed as words or clauses, and are both fuzzy numbers.

Examples :

i. Y_1 = Dinner's usually served at 9:00 pm.

Therefore, X = time of dinner service, and Z = <by 9:00pm, usually>

ii. Y_2 = This book has been a wonderful read.

Therefore, X = quality of book, and Z = <wonderful, certainly>

Understandably, *A* is context-dependent and explicitly mentioned while *B* is based on the perception of certainty presented by the statement. The value of *B* could be explicitly quoted in the statement (as in example i) or it could be implicit (as in example ii).

The ordered 3-tuple <*X*, *A*, *B*> is referred to as a Z-valuation. A Z-valuation is equivalent to an assignment statement *X* is <*A*, *B*>. As for example,

i. The Z-valuation of Y_1: <time of dinner service , by 9:00 pm, usually>

Implication: [time of dinner service] is (by 9:00 pm, usually)

ii. The Z-valuation of Y_2: <quality of book, wonderful, certainly>

Implication: [quality of book] is (wonderful, certainly).

A collection of Z-valuations is referred to as Z-information. The Z-information provides the impetus to a decision-making process.

Preliminary rules of computation using the Z-numbers:

i. For the purpose of computation, the values of *A* and *B* need to be precisiated through association with membership functions, μ_A, μ_B respectively.

ii. *X* and *A* together define a random event in *R*, and the probability of this event, *p*, may be expressed as:

$$p = \int_R \mu_A(u)\, p_X(u)\, du \tag{1}$$

where, *u* is a real valued generic value of *X* and p_X is the underlying (hidden) probability density of *X*.

iii. The Z-valuation <*X, A, B*> is viewed as a generalized constraint on *X*, and is defined by:

Probability (*X* is *A*) is *B*,

$$\text{or, } p = \int_R \mu_A(u)\, p_X(u)\, du \text{ is } B \tag{2}$$

iv. Equation 2 is mathematically equivalent to the expression:

$$\mu_B\left(\int_R \mu_A(u)\, p_X(u)\, du\right) \tag{3}$$

Subject to,

$$\int_R p_X(u)\, du = 1 \tag{4}$$

v. Computation using the Z-numbers is based on the Principle of Extension. As for example, considering a problem statement of the form:

"It is likely that the desert is good. What is the probability that it is not?"

Let,

X = desert, *A* = good, *B* = likely, *C* = not good, *D* = degree of certainty;

μ_A, μ_B, μ_C, μ_D are the membership functions associated with *A, B, C* and *D* respectively;

p_X is the underlying (hidden) probability density of *X*;

u is a real valued generic value of *X*.

Therefore, we have, (*X* is *A*) is *B*, and

We need to evaluate: (X is C) is ?D.

Thus, using the Principle of Extension and Equations 2, 3 and 4:

$$\frac{<X,A,B>}{<X,C,?D>}=\frac{\mu_B(\int_R \mu_A(u)p_X(u)du)}{(\int_R \mu_C(u)p_X(u)du) is ?D}$$

$$\text{or, } \mu_D(w)=\sup_{p_X} \mu_B(\int_R \mu_A(u)p_X(u)du) \tag{5}$$

Subject to,

$$w=(\int_R \mu_C(u)p_X(u)du) \text{ and}$$
$$\int_R p_X(u)du=1 \tag{6}$$

3.2 Probable Contributions of the Z-number to CWW

The definition of the Z-number subtly identifies the following as the contributions of the concept to the arena of CWW:

i. The Z-number serves as a model for the precisiation of natural language statements – unifying CWW and NLP.
ii. The Z-number summarizes the perception of a single simple sentence. If complex or compound sentences be deconstructed to their simple sentence components, the Z-number of each of the individual sentences can be evaluated to receive the Z-information for the entire sentence.
iii. The Z-information summarizes the perception of a group of statements.
iv. The parameters of Z-numbers help in the identification of the context of discourse of the sentence under observation – thus allowing sentences to be grouped into context-sensitive granules.
v. By virtue of points i though iv, the Z-number is visibly in agreement with the intuitive process of comprehension and reasoning in human beings.
vi. The Z-number could be used to extract 'knowledge' from a sentence.
vii. The parameters of the Z-numbers are context-independent.
viii. Translation from the Z-numbers to simple sentences is straightforward.

3.3 Algorithm for CWW following the Z-number Technique

Drawing inspiration from the possible areas of contribution of the Z-number to CWW, we aspire to design CWW-based systems that utilize the Z-number technique. It is thus that we have developed the following algorithm (Algorithm 1) for CWW based on the Z-number methodology. The algorithm takes natural language sentences as input and results in the natural language response to them as well.

Algorithm 1.

Input: Natural language sentence (I)

Output: Natural language response (O) to I

Assumptions:

i. The system is capable of distinguishing between relevant and irrelevant sentences
ii. The system comprehends the total perception of a complex or a compound sentence (Y) by –
    a. Extraction of the simple sentence components of Y
    b. Comprehension of the meaning of each of these simple sentence components

c. Combining these component perceptions with respect to the connectives in *Y*

[Assumption ii. follows the natural process of cognition in human beings]

Steps:

1. If *I* is irrelevant
   Then
   Goto step 10
   Else
   Goto step 2
2. If *I* is a simple sentence
   Then
   Goto step 3
   Else
   i. Extract the simple sentence component set (I') of *I*
   ii. Repeat steps 3 through 4 for each sentence in I'
   iii. Goto step 5.
3. Extract the values of *X*, *A* and *B* in the sentence to evaluate the Z-valuation (Z_I)
4. Convert Z_I into equivalent logical expression (Z_E)
5. Combine all Z_E to the logical expression (*E*) guided by the connectives in *I*
6. Convert *E* to the equivalent mathematical expression (*M*)
7. Evaluate *M* to receive a set of Z-valuations (Z_O) in response
8. Translate Z_O into simple sentences (*S*)
9. If step 8 results in more than one simple sentence
   Then
   If some or all the sentences in *S* can be compiled into a single sentence
   Then
   a. Assimilate all compatible simple sentences into a single complex or a compound sentence (S')
   b. If S' does not include all the sentences in *S*
      Then
      b.1. $S'' = S - S'$
      b.2. $O = S' \cup S''$
      Else
      $O = S'$
   c. Goto step 10
   Else
   a. $O = S$
   b. Goto step 10
   Else
   i. $O = S$
   ii. Goto step 10
10. Stop

4. Experiment

In this section, we describe an experiment where we strive to simulate the natural CWW by human beings. The machine tries to model the intuitive process of reasoning involved in the selection of a book at a bookstore.

Assuming that a human reader has a genre and content preference, the process of book selection intuitively follows the algorithm (Algorithm 2) outlined in Section 4.1.

Algorithm 2 is then extended to include elements of Z-number based CWW, as is described in Algorithm 1, so as to simulate the book selection process in a machine.

4.1 The Intuitive Algorithm Underlying the Process of 'Selection of a Book'

Algorithm 2.

Input:

i. The summary set (S) of a set of n ($n \geq 1$) books (B), where $S = \{S_1, S_2, \ldots, S_n\}$, $B = \{B_1, B_2, \ldots, B_n\}$ and S_i is the summary of B_i – extracted from the book jacket
ii. Event set (E) of the reader's expectation of the literary contents of the book to be chosen
iii. Lower threshold (T) indicating the minimum number of events, out of E, that need to exist in the S_i of the selected B_i
iv. Additional selection criteria (C) with respect to the external characteristics of the book required.

Output: Book selected

Steps:

1. For every $B_i \in B$
   Repeat steps 2 though 4
2. Read S_i
3. Extract event set E_S in S_i
4. If $|E \cap E_S| \geq T$
   Then
   Assign B_i a membership of selection (> 0)
   Else
   Membership of selection of $B_i = 0$
5. Sort B_i in the descending order of memberships of selection
6. If every book is assigned a unique membership of selection
   Then
   Select B_i with the highest membership as the book to be purchased
   Else
   i. Select all B_i with the highest membership
   ii. Identify the B_i that satisfies most of C
   iii. This B_i is selected as the book to be purchased
7. Stop

Note:

i. Each event in the event set (E) is implied by 'keywords' – the basis for the natural CWW by the reader.

ii. Besides the summary of the book, the following are natural contributing factors to the selection process and are elements of C :
   a. The book being read or present in the possession of the reader
   b. Cost of the book
   c. Condition of the book – brittleness, ink marks, crumpled pages, pages missing
   d. Presence and quality of pictures
   e. Font type and font size
   f. Rarity of the book

iii. In this experiment, the selection process is entirely platonic – devoid of any emotions occurring in real-life situations that lead to contemplation of an increase in budget or to the selection of a book quite different from that planned.

4.2 Simulation of the Process of 'Selection of a Book' Using the Z-number Technique

A. Intersection of Perceptions Conveyed by the Event-set (*E*) of the Reader's Expectations and the Event-set (E_s) Expressed in the Summary of the Book under Observation

An intersection of the perceptions conveyed by E and E_s should practically imply the extraction of the common meaning conveyed by these perceptions. It is thus that we define the perception-intersection operator $(\cap_p)$ as follows:

Let,

E and E_s consist of Z-valuations of the sentences expressing the reader's expectations and the summary of the book, respectively; and

$E' = \langle X_1, A_1, B_1 \rangle$ and $E_S' = \langle X_2, A_2, B_2 \rangle$ where $E' \in E$ and $E_S' \in E_S$

Then,

$$(E' \cap_p E_S') = \langle X_1, A_1, B_2 \rangle \text{ if } (A_1 \cap A_2) \neq \varnothing \quad (7)$$

X_1 and A_1 respectively represent a requirement and the corresponding expected value, while B_2 reflects the level of certainty with which the requirement is satisfied by the book under consideration.

The perception-intersection operator defined above could come of use in scenarios where it is imperative to verify the certainty with which a current situation satisfies a given requirement.

B. The Algorithm for the Process of 'Selection of a Book' Using the Z-number Technique

Combining elements of Algorithms 1 and 2, the algorithm for the book selection process using the Z-number Technique is outlined as follows:

Algorithm 3.

Input:

i. The summary set (S) of a set of n ($n \geq 1$) books (B), where $S = \{S_1, S_2, \ldots, S_n\}$, $B = \{B_1, B_2, \ldots, B_n\}$ and S_i is the summary of B_i – extracted from the book jacket

ii. Event set (E) – as Z-valuations – of the reader's expectation of the literary contents of the book to be chosen

iii. Lower threshold (T) indicating the minimum number of events, out of E, that need to exist in the S_i of the selected B_i

iv. Additional selection criteria (C) with respect to the external characteristics of the book required.

Output: Book selected

Assumptions:

i. A sentence is considered relevant if it contains at least one keyword from E

ii. Any word (W) in the summary that is a synonym of a keyword (W') in E, is treated as W'

iii. The system comprehends the total perception of a complex or a compound sentence (Y) by –

  a. Extraction of the simple sentence components of Y

  b. Comprehension of the meaning of each of these simple sentence components

  c. Combining these component perceptions with respect to the connectives in Y

iv. The memberships of selection are assigned on the basis of the principle of extension, explained in Section 3.1

v. The reader does not read the summary of a book he/she has read or possesses.

Steps:

1. For every B_i in B
   Repeat steps 2 though 10
2. Initialize $E_S = \varnothing$ [E_S = Event set for the current B_i]
3. For every sentence (I) in S_i
   Repeat steps 4 through 7
4. If I is irrelevant
   Then
      i. Discard I
      ii. Goto next I
   Else
      Goto step 5
5. If I is a simple sentence
   Then
      Goto step 6
   Else
      i. Extract the simple sentence component set (I') of I
      ii. Repeat steps 6 through 7 for each sentence in I'
      iii. Goto step 8.
6. Extract the values of X, A and B in the sentence to evaluate the Z-valuation (Z_I)
7. Combine all Z_I to the logical expression (E') guided by the connectives in I
8. $E_S = E_S \cup E'$
9. If $| E \cap_p E_S | \geq T$
   Then
      i. Convert ($E \cap_p E_S$) to the logical expression ($\tilde{E}$)
      ii. Convert E and $\tilde{E}$ to the mathematical expressions M and $\tilde{M}$ respectively.
      iii. Evaluate the membership of selection of B_i by applying the principle of extension on M and $\tilde{M}$

Else

Membership of selection of $B_i = 0$

10. Sort B_i in the descending order of memberships of selection
11. If every book is assigned a unique membership of selection

Then

Select B_i with the highest membership as the book to be purchased

Else

i. Select all B_i with the highest membership
ii. Identify the B_i that satisfies most of C
iii. This B_i is selected as the book to be purchased

12. Stop

C. Simulation of Algorithm 3

Assumptions:

i. The machine (*Mc*) is capable of annotating the words in a given text sample into the correct parts of speech and is capable of identifying the type of the sentence under consideration.

ii. *Mc* has read thirty works of fiction in the ‘Mystery’ genre.

iii. *Mc*’s vocabulary consists of one hundred and sixty five keywords (X). Each of these words is assigned a probability of occurrence (p_x), based on the number of books the words are found in.

The words in *Mc*’s vocabulary are: abduct, accomplice, advocate, agent, alibi, allegation, ammunition, anonymous, arms, assistant, awkward, baffle, blood, blunder, bury, case, catch, chief, chilling consequence, clue, cold-blooded, conspiracy, constable, convict, corpse, crime, criminologist, crooked, curious, danger, death, deceive, deduce, desperate, detective, discover, doctor, drug, duplicate, eavesdrop, enemy, evidence, evil, exhume, fake, fatal, fear, figure-out, find-out, fingerprint, follow, forbidden, forget, foul play, gang, gore, graveyard, gray cells, guilty, headquarters, hidden, hoax, homicide, how, illegal, illegitimate, illicit, impersonate, ingenuity, innocence, inquest, inspector, instinct, intrigue, investigate, jewels, judge, juvenile, kidnap, kill, lawyer, letters, locate, loot, macabre, Marple, mask, missing, mistake, motive, murder, mystery, nab, notorious, overhear, peculiar, plan, plot, plunder, Poirot, poison, police, post mortem, practical joke, prison, problem, proof, prosecution, psychology, puzzle, quarrel, question, racket, ransack, ransom, realize, red-handed, remember, remorse, remorseless, replicate, revenge, robber, sabotage, scandal, scheme, Scotland Yard, secret, sentence, shocking, shoot, sinister, soldier, solve, spy, stab, stolen, strange, suicide, superintendent, surprise, suspect, suspicious, symbols, terror, thief, tragic, trap, trial, trouble, underground, unknown, vengeance, verdict, victim, vile, violence, warning, weapons, what, when, where, who, whom, why, witness.

Mc is aware of synonyms like: (kill, murder), (problem, puzzle), (deduce, find-out, figure-out), (anonymous, unknown), (suspicious, curious), (verdict, sentence), (plot, scheme) etc.

Mc is aware of the polysemous/homonymous nature and different forms of certain words like: (puzzle (noun, verb), puzzled, puzzling), (judge (noun, verb)), (mystery, mysterious), (murder (noun, verb), murderer) etc.

iv. The words in *X* are subdivided into categories like: mystery_category, detective_name, events, verdicts, and so on where each such generic value is mapped to a real number – following the definition of the Z-number.

v. The words are further clustered into semantic nets or groups of words that occur together or are semantically linked e.g. Murder_net <murder, motive, quarrel, post mortem, police, exhume, fatal, clue, revenge> etc.

Inputs:

i. *Mc* wants to buy a new book with the requirements as follows:

Table 1. Requirement Set of M

Requirements – Event Set (*E*)	
Natural Language Statements	**Z-valuations**
The book needs to be a Mystery story	Z_1 = <Book Genre, mystery, certainly> A = mystery, u_1 = book genre
Preferably a murder mystery solved by Miss Marple	Z_{21} = <Mystery category, murder, ideally> A = murder, u_2 = mystery category Z_{22} = <Detective, <Marple, ideally> A = Marple, u_3 = detective
A mystery involving a grand robbery and solved by Marple or Poirot would be okay	Z_{31} = <Mystery category, robbery, possibly> A = robbery, u_2 = mystery category Z_{32} = <Detective, Marple, probably> A = Marple, u_3 = detective Z_{33} = <Detective, Poirot, probably> A = Poirot, u_3 = detective
Requirements – Others (*C*)	
Price ≤ Rs. 300/-	
New – no crumpled pages or ink marks	
M should not have already read the book	

Note:

a. Parameters *A* and *B* are quintessentially maintained and processed as fuzzy set types specific to the system.

b. [11] professes the Interval-Type2 Fuzzy Set being the most appropriate model of word perceptions.

Thus, *E* can be summarized by the logical expression:

$$[Z_1 \wedge ((Z_{21} \wedge Z_{22}) \vee (Z_{31} \wedge (Z_{32} \vee Z_{33})))] \tag{8}$$

And following Equations 2, 3 and 4, as discussed in Section 3.1, Equation 8 may be rewritten as:

$$M = \mu_{certainly}(\int_R \mu_{mystery}(u_1)\, p_X(u_1)du_1) \wedge [\{\mu_{ideally}(\int_R \mu_{murder}(u_2)\, p_X(u_2)du_2) \wedge \mu_{ideally}(\int_R \mu_{Marple}(u_3)\, p_X(u_3)du_3)\} \vee [\mu_{possibly}(\int_R \mu_{robbery}(u_2)\, p_X(u_2)du_2) \wedge \{\mu_{probably}(\int_R \mu_{Marple}(u_3)\, p_X(u_3)du_3) \vee \mu_{probably}(\int_R \mu_{Poirot}(u_3)\, p_X(u_3)du_3)\}]] \quad (9)$$

where,

$$\int_R p_X(u_1)du_1 = 1, \int_R p_X(u_2)du_2 = 1, \int_R p_X(u_3)du_3 = 1$$

ii. $T = 2$

iii. The summaries of some new books in the store are as follows:

a. **Summary:** Lymstock is a town with more than its share of secrets – a town where even a sudden outbreak of anonymous hate-mail causes only a minor stir. But all that changes when one of its recipients, Mrs. Symmington, commits suicide. Her final note said, “I can’t go on.” Only Miss Marple questions the coroner’s verdict of suicide. Was this work of a poison-pen? Or of a poisoner? – The Moving Finger (Agatha Christie)

   Other Properties: Price = Rs. 150/-; New book; Not read

b. **Summary:** “The curious case of the Maiden Eggesford Horror”. When the doctor advises Bertie to live the quiet life, he and Jeeves head for the pure air and peace of Maiden Eggesford. However, they hadn’t reckoned on Bertie’s irrepressible but decidedly scheming Aunt Dahlia, around whom an imbroglio of impressive proportions develops involving The Cat Which Kept Popping Up When Least Expected. As Berties observes, whatever aunts are, they are not gentlemen. – Aunts Aren’t Gentlemen (P. G. Wodehouse)

   Other Properties: Price = Rs. 250/-; New book; Not read

c. **Summary:** Gerry Wade had proved himself to be a champion sleeper; so the other house guests decided to play a practical joke on him. Eight alarm clocks were set to go off, starting at 6:30 a.m. But when morning arrived, one clock was missing and the prank had backfired with tragic consequences. Gerry never woke up. Was he murdered? – The Seven Dials Mystery (Agatha Christie)

   Other Properties: Price = Rs. 150/-; New book; Not read

Execution:

Table 2. Summarization of Book 1

Relevant sentences in the summary	Simple sentence components	Z-valuations of the simple sentence components
Lymstock is a town with more than its share of secrets – a town where even a sudden outbreak of anonymous hate-mail causes only a minor stir.	1. Lymstock is a town with secrets. 2. There is a sudden outbreak of anonymous hate-mail.	Z_{11} = <Location, Lymstock, supposedly> Z_{12} = <Location property, has secrets, supposedly> Z_{13} = <Event, anonymous letters, probably> [hate-mail = letters]

But all that changes when one of its recipients, Mrs. Symmington, commits suicide.	1. Recipient Mrs. Symmington commits suicide.	Z_2 = <Letter event, recipient commits suicide, probably>
Only Miss Marple questions the coroner's verdict of suicide.	1. Coroner's verdict is suicide. 2. Miss Marple questions verdict.	Z_{31} = <Coroner verdict, suicide, probably> Z_{32} = <Verdict event. Marple questions, probably>
Was this work of a poison-pen?	- Is a simple sentence -	Z_4 = <Suspect, poison-pen, expectedly>
Or of a poisoner?	- Is a simple sentence -	Z_5 = <Suspect, murderer, expectedly> [poisoner = murderer]

Therefore,

$$E_{S1} = (Z_{11} \wedge Z_{12} \wedge Z_{13}) \wedge Z_2 \wedge (Z_{31} \wedge Z_{32}) \wedge (Z_4 \vee Z_5);$$

$|E \cap_p E_{S1}| = 3 > T$ [By virtue of Z_{31} and Z_5 in the summary of book1; and the fact that quite a large number of words in the summary of book1 fall under the vocabulary of *Mc*, the book certainly pertains to the genre 'Mystery'];

$(E \cap_p E_{S1})$ = [<Book genre, mystery, certainly> ∧ <Detective, Marple, probably> ∧ <Mystery category, murder, expectedly>] (Using Equation 7) and the corresponding mathematical expression (using parameters of Equation 9) is:

$$\tilde{M} = \mu_{certainly}\left(\int_R \mu_{mystery}(u_1)\, p_X(u_1)\, du_1\right) \wedge \left(\mu_{probably}\left(\int_R \mu_{Marple}(u_3)\, p_X(u_3)\, du_3\right) \wedge \mu_{expectedly}\left(\int_R \mu_{murder}(u_2)\, p_X(u_2)\, du_2\right)\right) \quad (10)$$

The membership of selection for book1 is evaluated on the basis of the application of the principle of extension to Equations 9 and 10; and considering the degree of overlap, the membership of selection should ideally approximate 1.

Table 3. Summarization of Book 2

Relevant sentences in the summary	**Simple sentence components**	**Z-valuations of the simple sentence components**
The curious case of the Maiden Eggesford Horror.	- Is a simple sentence -	Z_1 = <Case, Maiden Eggesford Horror, supposedly>
However, they hadn't reckoned on Bertie's irrepressible but decidedly scheming Aunt Dahlia, around whom an imbroglio of impressive proportions develops involving The Cat Which Kept Popping Up When Least Expected.	1. Aunt Dahlia is decidedly scheming	Z_{21} = <Character, Aunt Dahlia, certainly> Z_{22} = <Character nature, scheming, decisively>

As Bertie observes, whatever aunts are, they are not gentlemen.	1. Bertie observes that aunts aren't gentlemen.	Z_{31} = <Character, Bertie, certainly> Z_{32} = <Character action, observation, decisively>

Therefore,

$$E_{S2} = (Z_1 \wedge Z_{21} \wedge Z_{22} \wedge Z_{31} \wedge Z_{32}) \;;$$

$|E \cap_p E_{S2}| = 1 < T$ [By virtue of the fact that some words in the summary of book2 fall under the vocabulary of *Mc*, the book 'probably' pertains to the genre 'Mystery'].

Thus, book2 is assigned a membership of selection = 0

Table 4. Summarization of Book 3

Relevant sentences in the summary	Simple sentence components	Z-valuation of the simple sentence components
Gerry wade had proved himself to be a champion sleeper; so the other house guests decided to play a practical joke on him.	1. The house guests played a practical joke on Gerry.	Z_1 = <Event, practical joke, supposedly>
But when morning arrived, one clock was missing and the prank had backfired with tragic consequences	1. One clock was missing. 2. Prank had backfired. 3. Consequences were tragic.	Z_{21} = <Event, clock missing, supposedly> Z_{22} = <Event, practical joke, supposedly>[prank = practical joke] Z_{23} = <Practical joke event, backfire, supposedly> Z_{24} = <Practical joke event, tragic consequence, supposedly>
Was he murdered?	- Is a simple sentence -	Z_3 = <Event, murder, expectedly>

Therefore,

$$E_{S3} = (Z_1 \wedge Z_{21} \wedge Z_{22} \wedge Z_{23} \wedge Z_{24} \wedge Z_3);$$

$|E \cap_p E_{S3}| = 2 = T$ [By virtue of Z_3 in the summary of book3; and the fact that quite a large number of words in the summary of book1 fall under the vocabulary of *Mc*, the book certainly pertains to the genre 'Mystery'];

$(E \cap_p E_{S3})$ = [<Book genre, mystery, certainly> ∧ <Mystery category, murder, expectedly>] (Using Equation 7), and the corresponding mathematical expression (using parameters of Equation 9) is:

$$\tilde{M} = \mu_{certainly}\left(\int_R \mu_{mystery}(u_1)\, p_X(u_1) du_1\right) \wedge \mu_{expectedly}\left(\int_R \mu_{murder}(u_2)\, p_X(u_2) du_2\right) \qquad (11)$$

The membership of selection for book3 is evaluated on the basis of the application of the principle of extension to Equations 9 and 11, and considering the degree of overlap, the membership of selection should lie in the range (0, 1) and should be less than that of Book1.

Thus, on the basis of the interpretations of Equations 10 and 11, *Mc* evidently selects book1 – “The Moving Finger” by Agatha Christie. This decision by *Mc* coincides with the judgment a human being would make, given the scenario.

D. Observations

The experiment leads to the following illuminating observations:

i. The Z-number methodology blends in seamlessly with the intuitive algorithm followed for the example studied.

ii. CWW based on the Z-number calls for:

   a. A powerful Part-Of-Speech (POS) tagger – to identify the parts of speech of the words in the sentences – an insight into the probable values of *X*, *A* and *B* indicate noun phrases, adjectives and adverbs respectively.

   b. Resolution of the dependencies between the nouns and the corresponding pronouns used – such that the sentences on a common topic might be identified and processed accordingly.

   c. Methods to reduce complex or compound sentences to their simple sentence counterparts and vive versa.

   d. A robust model for the representation of the perceptions of the values of *A* and ***B*.**

   e. Compilation of a comprehensive context-sensitive text corpus – such that the probable generic values for the parameter *X* can be listed and mapped to real numbers, and the probable values of the parameter *A* be identified and assigned membership values and probability distributions as required.

   f. An algorithm to identify keywords and relevant sentences.

   g. An algorithm to extract implicitly specified values of the parameter *B*.
   A possible solution could be the use of default values with respect to the type of the sentence, as is shown in Table 5. (All examples in this article use this technique.)

Table 5. Suggested default values for B, with respect to the type of sentence.

Sentence Type	**Remarks**	**Default Value for *B***
Declarative		Probably or Supposedly
Exclamatory	Conveys explicit emotion	Certainly
Imperative	Expect a definite course of action	Definitely or Decisively
Interrogative	Expects an answer	Expectedly

Note: In a natural conversation scenario – the behavioural parameters (general attitude) of the participants in a discourse influence the value of *B*.

h. An algorithm to identify and process polysemous and homonymous words, capitonyms and synonyms.

i. Methods to identify new words, map the meanings of these words to synonyms existing in the system vocabulary and to learn to use them as well.

j. Schemes to identify and process new sub-contexts under the context of discourse.

k. Schemes to translate Z-valuations to logical expressions, mathematical expressions and natural language statements.

l. Well defined operators and rules of computation – as is guided by the context and the word-perception models assumed.

Each of these aforementioned points (from ii.a though ii.l) represent the basic challenges underlying the implementation of the Z-numbers for CWW.

iii. The system should respond within the average human response time (150 – 300 ms).

5. Conclusion

This article is an elucidation on our study of the Z-number approach to CWW. The Z-number not only aspires to provide a framework for the precisiation of the meaning of a natural language statement, but includes parameters that permit extraction of the information and the uncertainty of the information conveyed by the sentence. The Z-numbers predictably have a radical role to play in the realm of CWW and NLP. The probable contributions of the Z-number to CWW have been clearly outlined in the article.

Besides the strengths, we present here algorithms that illustrate the seamless merge of intuition with Z-number elements to simulate a real-life CWW scenario. These algorithms illuminate the need for intensive research on a fair number of issues that need to be resolved prior to the actual implementation of the Z-numbers. Some of the prime areas that need delving into are models of word perceptions, algorithms for the extraction of the Z-number parameters from natural language statements, well-defined rules of computation, and algorithms to convert Z-numbers to logical/mathematical expressions and natural language statements. The article also defines an operator for the intersection of perceptions.

The Z-number is indeed an intriguing field of study in the arena of CWW and we are in the process of excavating the answers to the challenges highlighted in this article.

6. Acknowledgement

This project is being carried out under the guidance of Professor Sankar K. Pal. The work was supervised by Prof. Pal as a J.C. Bose Fellow of the Government of India.

References

[1] A. M. Turing, Computing machinery and intelligence, *Mind*, vol. 49, pp. 433–460, 1950.

[2] L. A. Zadeh, Intelligent systems revolution: Is it Real?, *http://www-bisc.cs.berkeley.edu/zadeh/papers*, 1998.

[3] L. A. Zadeh, Fuzzy logic = Computing with words, *IEEE Trans. Fuzzy Systems*, vol. 4, no. 2, pp. 103–111, 1996.

[4] L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. Syst. Man Cyberns.*, vol. 3, no. 1, pp. 28–44, 1973.

[5] L. A. Zadeh, Test-score semantics for natural languages and meaning representation via PRUF, *Empirical Semantics*, B. Reiger (ed), Brockmeyer, W.Germany, pp. 281–349, 1982.

[7] L. A. Zadeh, Precisiated Natural Language (PNL), *AI Magazine*, vol. 25, no. 3, pp. 74–91, 1994.

[6] L. A. Zadeh, A new direction in AI: Toward a computational theory of perceptions, *AI Magazine*, vol. 22, no. 1, pp. 73–84, 2001.

[8] L. A. Zadeh, A note on Z–Numbers, *Information Sciences*, vol. 181, pp. 2923–2932, 2011.

[9] J. Kacprzyk and S. Zadrozny, Computing with words is an implementable paradigm: Fuzzy queries, linguistic data summaries and natural language generation, *IEEE Trans. Fuzzy Systems*, vol. 18, no. 3, pp. 461–472, 2010.

[10] J. C. Nesfield, *English Grammar Series: Book IV: Idioms, grammar and synthesis*, Macmillan, London, 1956.

[11] J. M. Mendel and D. Wu, *Perceptual computing – Aiding people in making subjective judgments*, IEEE Press, N.J., 2010.