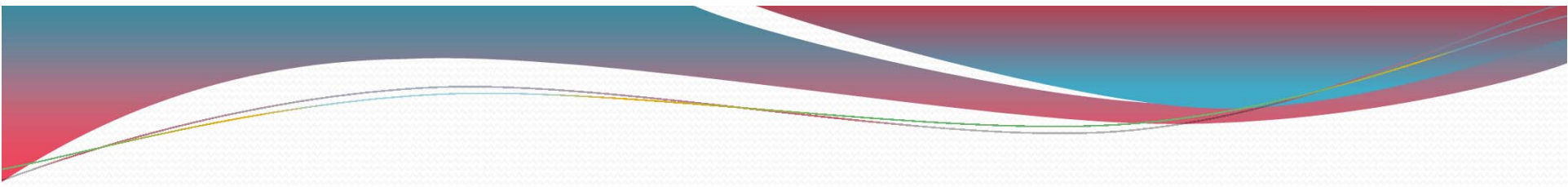


Tutorial

Fuzzy Logic, Probability, and Measurement: Similarities and Differences in Computing with Words (CWW)

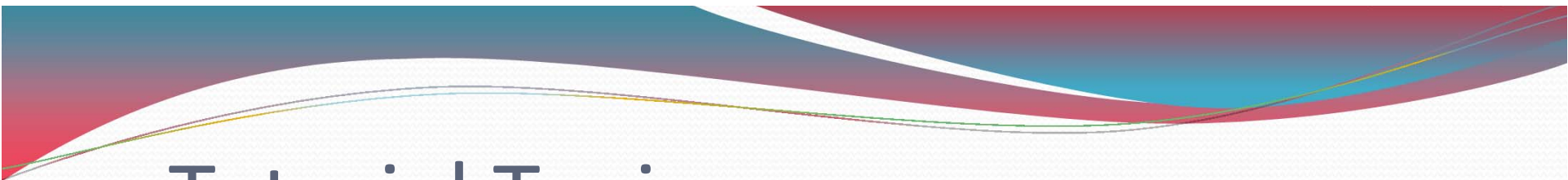
Boris Kovalerchuk
Dept. of Computer Science
Central Washington University, USA
borisk@cwu.edu

June 10, 2012, WCCI 2012, Brisbane
3 PM-5 PM



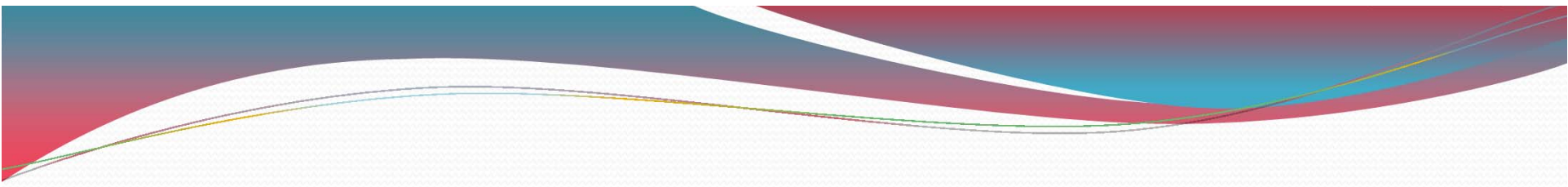
Agenda at glance

- Background
- Approaches
- Advantages and disadvantages
- Open questions for the future research
- Conclusion



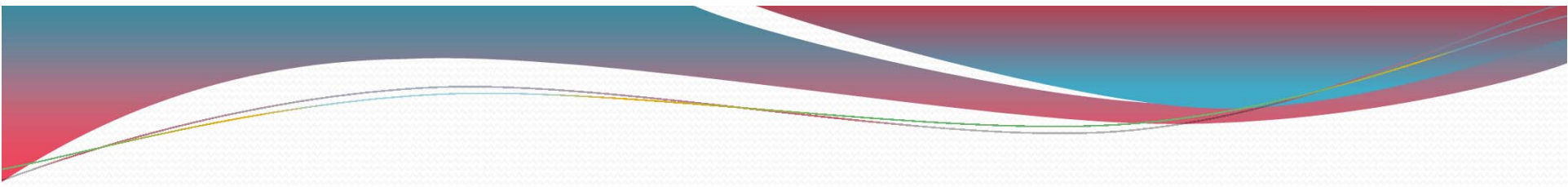
Tutorial Topics

- Links of *Fuzzy* and *Probability* concepts with the Representative *Measurement* theory from the mathematical psychology
- *Probabilistic Interpretation* of Fuzzy sets for CWW
- Interpretation of fuzzy sets and probabilities in a *linguistic context spaces*
- Interpretation of fuzzy sets with the concept of *irrational agents*



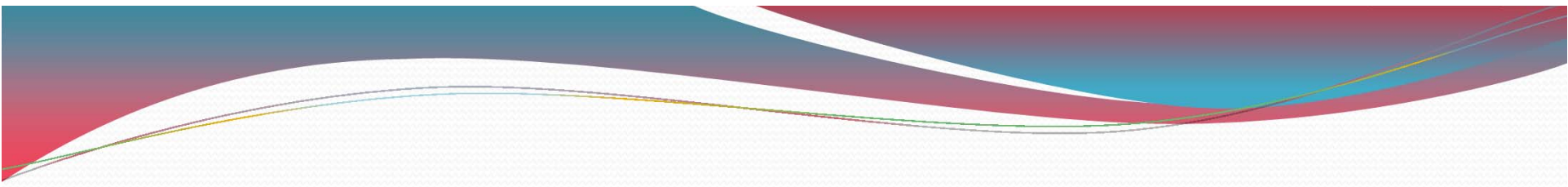
Agenda

- Topics
- Motivation
- Computing with words as “human logic” modeling



Tutorial Motivation

- **Fuzzy Logic, Probability, and Measurement for CWW**
- L. Zadeh initiated Computing with Words (CWW) studies with fuzzy logic and probability combined. He proposed a set of CWW test tasks and asked whether probability can solve them.
- CWW and mathematical principles for modeling **human logic under uncertainty**
 - comparing probabilistic, stochastic and fuzzy logic approaches.
- The most recent discussion on this topic started at the BISC group with a “naïve” question from a student: “What is the difference between fuzzy and probability?” that generated multiple answers and a very active discussion.
- Numerous debates can be found in the literature [e.g., Cheeseman; Dubois, Prade; Hisdal; Kovalerchuk,] and BISC 2009-2012 about relations between fuzzy logic and probability theory.



Panel on CWW

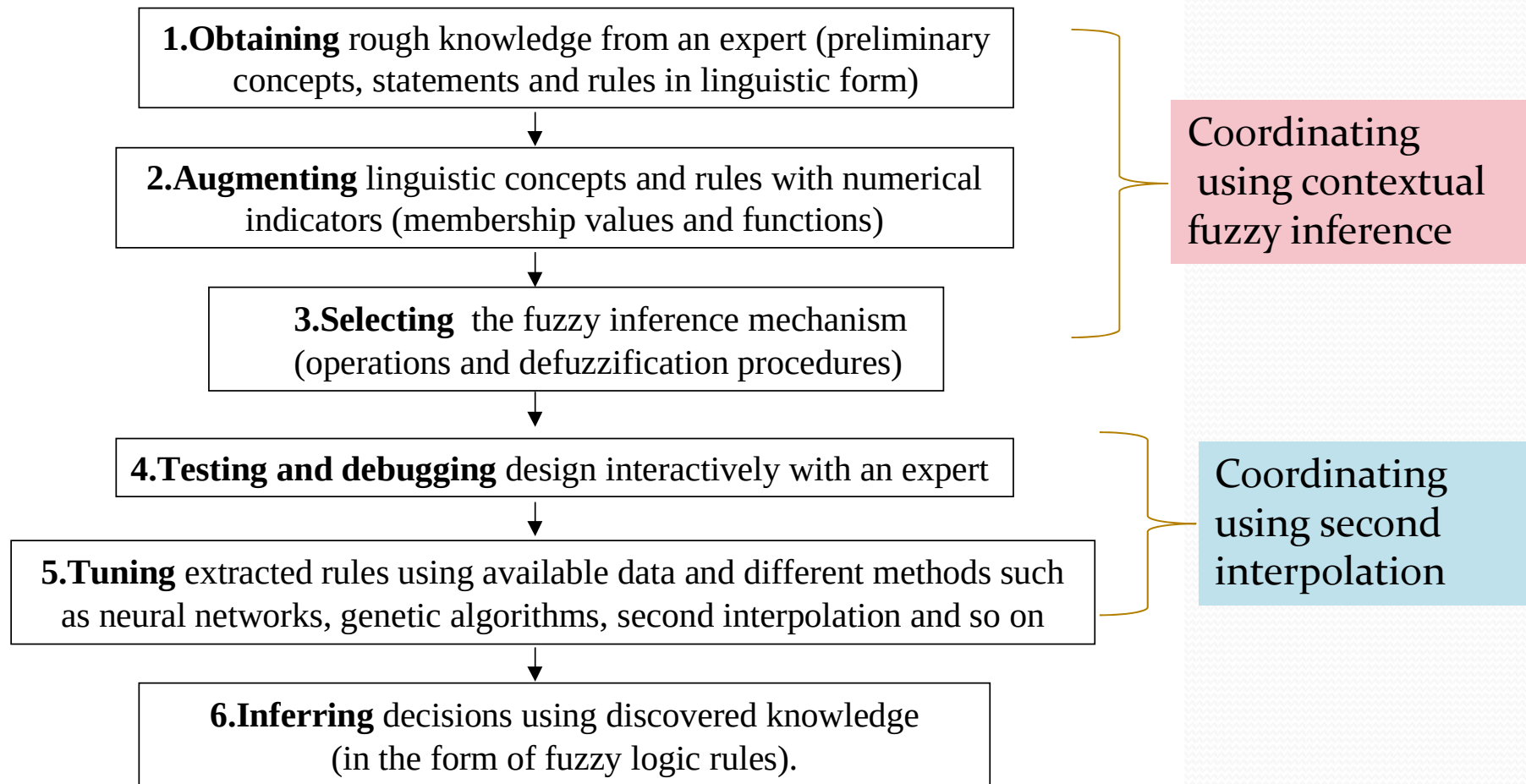
- This tutorial can be a starting “warming up” point for the round table on the same topic that the presenter and V. Kreinovich organized at WCCI 2012.
- The previous related round table on uncertainty modeling was at the World Conference on Soft Computing in May 2011 with invited panelists (Zadeh, Widrow, Kacprzyk, Kovalerchuk, Perlovsky)

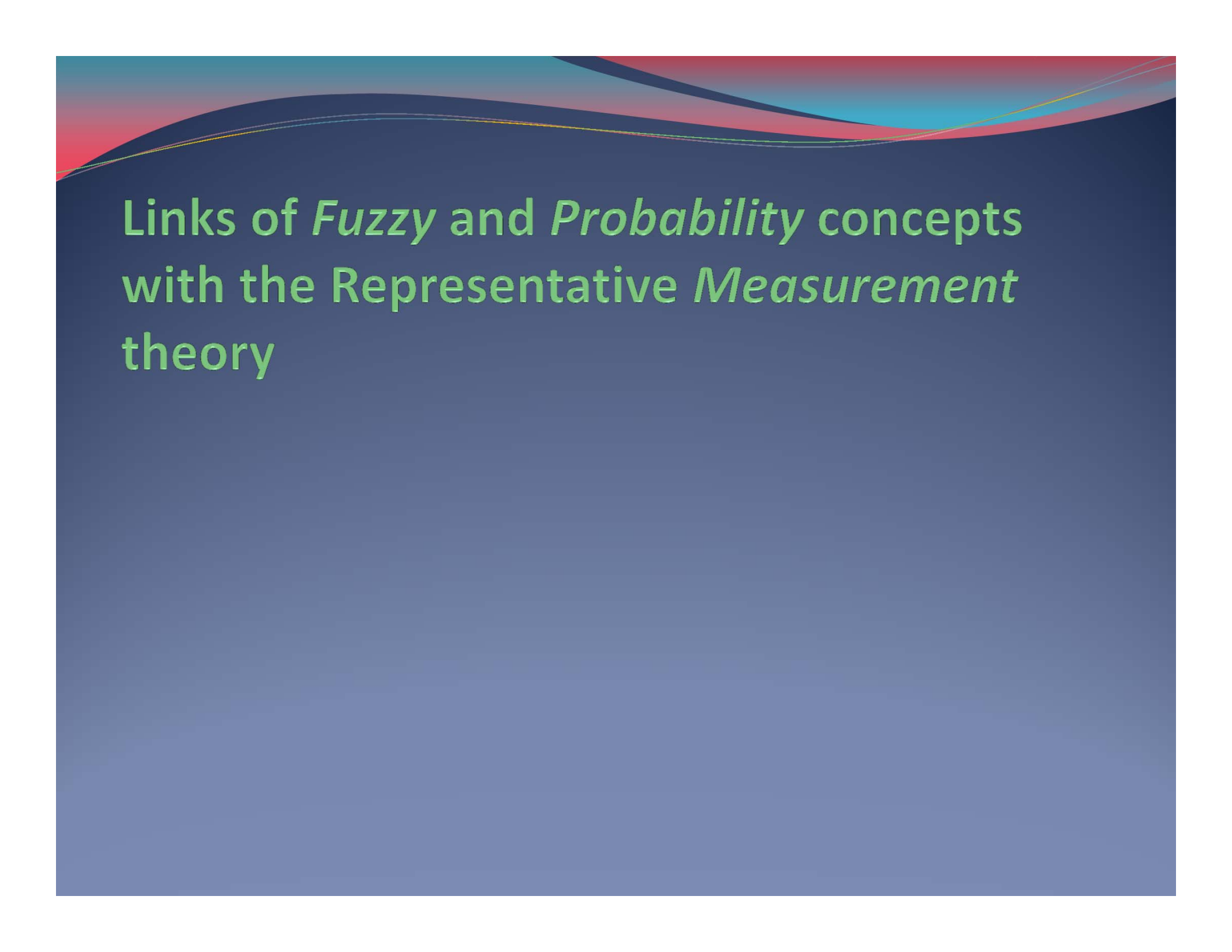
Quote: "All models are wrong, but some are useful."

Computing with words as “human logic” modeling

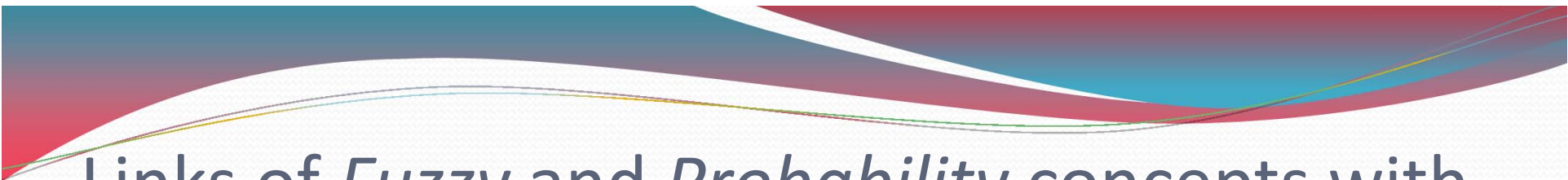
- Modern approaches to modeling “human logic” under uncertainty range from
 - classic Probability Theory to Fuzzy Logic, Dempster-Shafer theory, Rough Sets, Probabilistic Bayesian Networks, etc.
- All such approaches are oriented to somewhat **different contexts**.
 - However, the appropriate context for a given application often is not clearly formulated, and thus it is very difficult to (*a priori*) select one of the approaches in favor of another in many situations.

Knowledge discovery and Computing with Words based on Fuzzy Logic





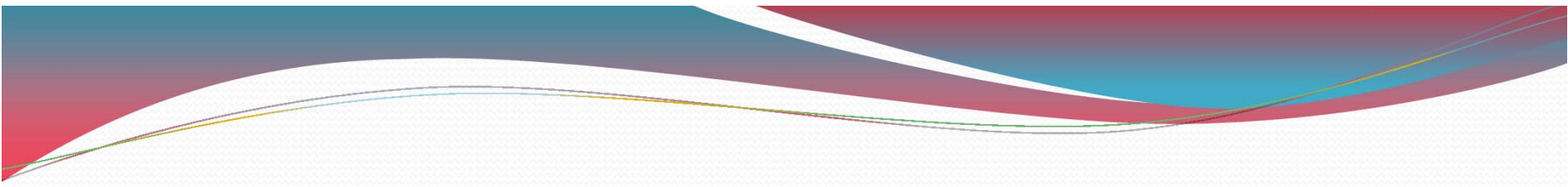
Links of *Fuzzy* and *Probability* concepts with the Representative *Measurement* theory



Links of *Fuzzy* and *Probability* concepts with the Representative *Measurement* theory

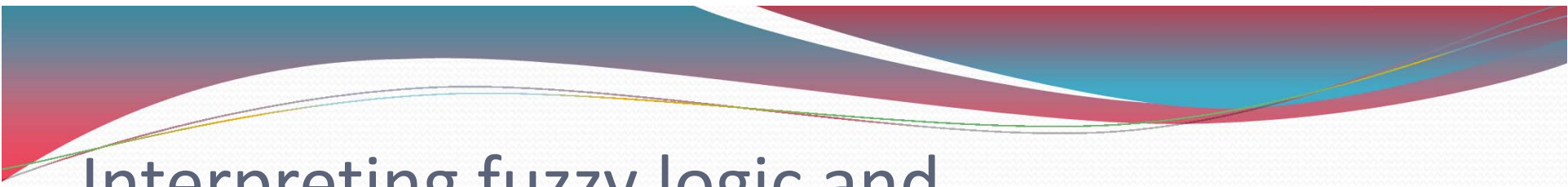
What are individual and joint areas for fuzzy logic and probability theory from the viewpoint of measurements?

- The viewpoint of measurements means that we do not start from given probabilities P and membership functions (MFs), but
 - we need to obtain P and MF as well as
 - to justify reasoning operations with P and MF.
- What is an appropriate way to obtain fuzzy membership function (MF) values, or in other words how to measure membership functions?
- If MF presents a subjective human judgment, how to distinguish MF from subjective probabilities?



How to get membership functions?

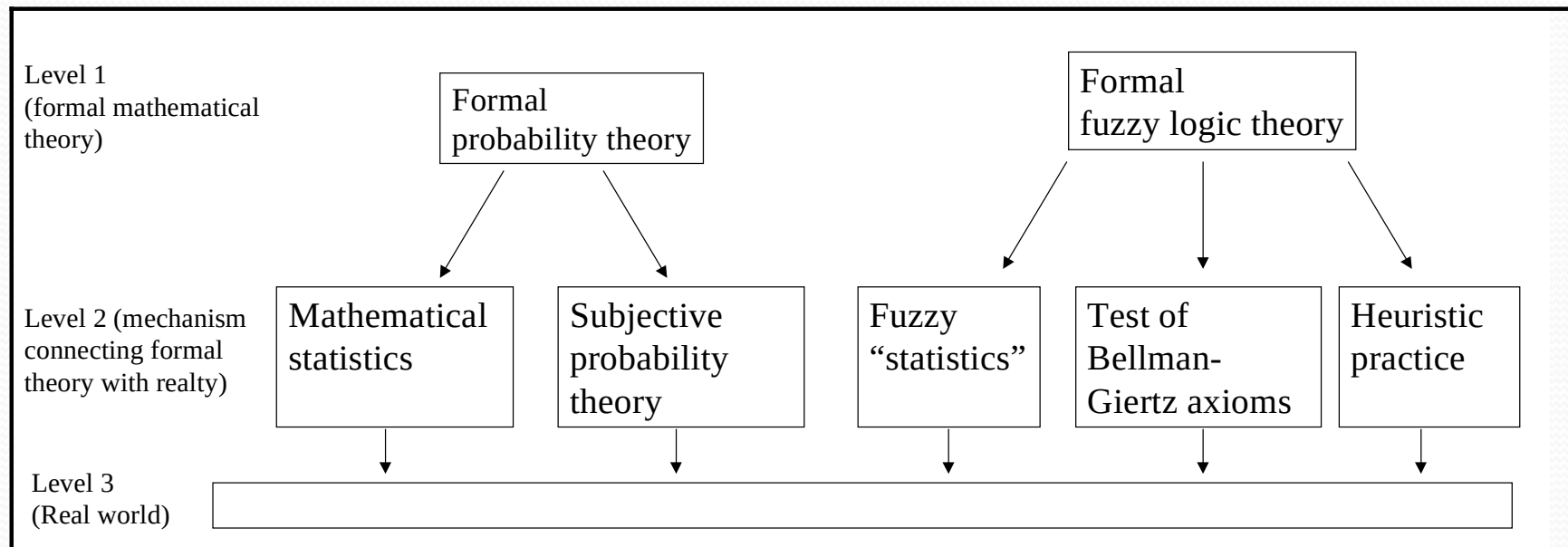
- Fuzzy membership functions (MF) often are produced (measured) by using frequencies which is a probabilistic way to get MFs.
- The user of these MFs should get an answer for the question: “Why should we use T-norms and T-conorms with these “probabilistic” MFs instead of probabilistic operations?”
- The same question is important from the theoretical viewpoint.



Interpreting fuzzy logic and probability theory

- **Interpreting mechanisms**
 - interpret theoretical concepts with real world entities.
- The major theoretical concept of probability theory to be interpreted is
 - the **concept of probability $p(x)$** and
- The major theoretical concept of fuzzy logic to be interpreted is
 - the **concept of membership function $m(x)$** .
- Both serve as **measures of uncertainty** of real-world entities.
- Interpreting mechanisms should also interpret **operations** with $p(x)$ and $m(x)$ as meaningful operations for real-world entities.

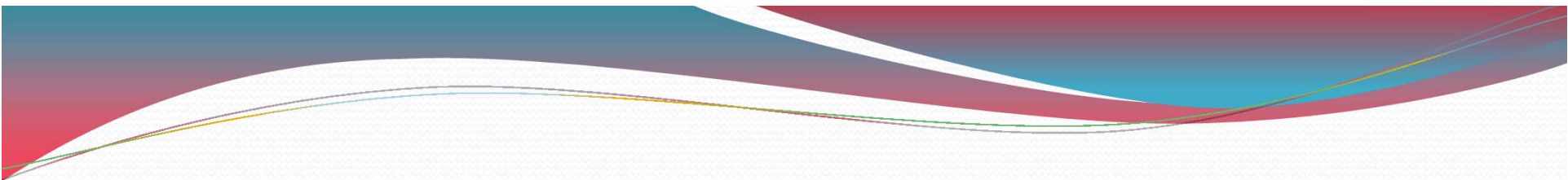
Acquisition of MFs and the difference between fuzzy logic and probability theory



Fuzzy Logic and Probability Theory:
Probabilistic Interpretations of Fuzzy
Sets for CWW

Comparison of stochastic and linguistic uncertainties

<i>Feature</i>	<i>Stochastic uncertainty</i>	<i>Linguistic uncertainty</i>
<i>Examples of statements</i>	Statement 1: "The probability of DJIA to be between 11000 and 11100 on June 1st is 0.7" Statement 3: "It is possible to put n passengers into Carole's car" [Cheeseman, 1985]	Statement 2: "We will probably have a successful financial year" [Von Altrock, 1997]. Statement 4: "It is possible that shareholders will be satisfied with the mutual fund performance"
<i>Clarity of event definition</i>	Clearly defined events: "DJIA will be between 11000 and 11100 on June 1st ", "put n passengers into Carole's car"	Not clearly defined events: "successful financial year", "satisfied shareholders of the mutual fund"
<i>Exactness of resulting probability</i>	High	Low
<i>Importance of background of a person making an evaluation</i>	Relatively low	High
<i>Type of concept</i>	Objective	Subjective
<i>Source of uncertainty</i>	Occurrence of event	Imprecision of human language



Probabilistic Interpretation of Fuzzy Sets

1. Interpretation of *Fuzzy* sets as *Random* Sets
2. Fuzzy sets and *Statistical Probability*
3. Fuzzy sets and *Kolmogorov's axioms* of the probability: the role of the mutual exclusion axiom
4. Interpretation of fuzzy sets with *subjective probabilities*
5. Interpretation of fuzzy sets and probabilities in a *linguistic context spaces*
6. Comparison of T-norms, T- conorms and probabilistic *operations* for fuzzy sets.
7. Comparison of fuzzy and probabilistic approaches for the application tasks with *extensive reasoning*



Reasoning in Fuzzy logic and Probability theory

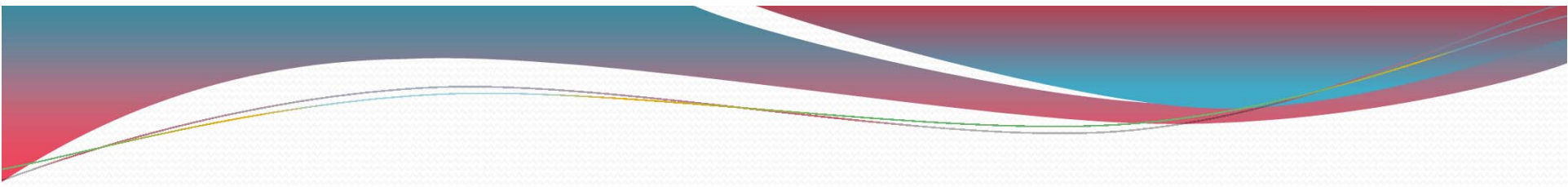
There are tasks where a subjective judgment is used directly and there are tasks where subjective judgments need to be combined, fused to infer a conclusion.

- How to select and justify the reasoning rules and operations when we need to combine judgments?
 - For instance, we have multiple T-norms and T-conorms used in fuzzy logic and probabilistic rules used in the probability theory, such as the Bayesian rule, union and intersection of probabilities.
- What is the justification to use the fuzzy logic operations when membership values are obtained by using a probabilistic technique such as frequencies or random sets?



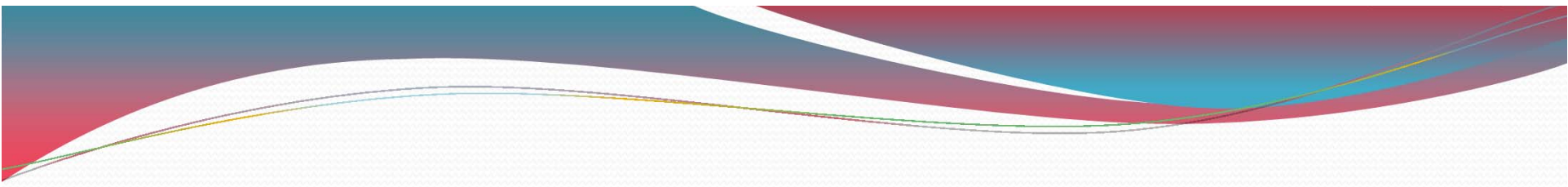
Scientific rigor as a guidance for fuzzy-probability comparison

- Chances to find a common ground in the relations between fuzzy and probability paradigms and their mutual benefits are higher through the consensus on the scientific rigor of the solutions than without it.
- The “pessimistic” stance is that due to the fact that Natural Language (NL) is inherently uncertain and the tasks in NL are ill-posed, the scientific rigor is out of reach.
- I am on the “optimistic” side that the scientific rigor in CWW is within reach.
- It can be done by the appropriate combination of concepts from fuzzy logic, probability theory, representative measurement theory, agent-bases approaches, and others.
- The advantage of the “optimistic” stance is that it gives a solid ground for the further progress of the field in combination with the already existing strong applications.



Scientific Rigor (SR)

- Example of SR level:
 - It is unreasonable to expect that Ancient Rome engineers would design bridges with modern rigor.
 - It would be also unreasonable for the modern civil engineers to decrease the SR to what was appropriate for the ancient engineers.
 - The modern bridge design is much more rigorous. We are more certain that the bridge will support the expected load. It is not a full determinism. It may suffer from uncertainty of the weather.
- The concept of SF is a *dynamic* one.
 - People in the next 2000 years may consider the modern SR level of bridge design as very low similar to how we see the level in the Ancient Rome.
- In general SR is *task specific*. SR depends on:
 - task itself, task time horizon, and current state of the art of the field that includes the level of the determinism-uncertainty of its theory.



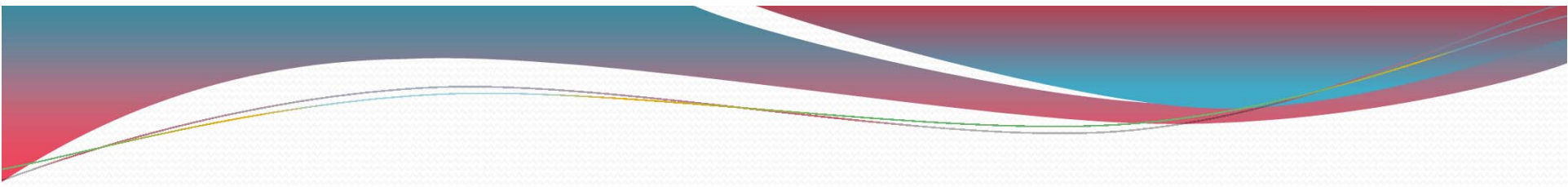
Scientific Rigor for FL and PT

- For the fuzzy logic and the probability theory the main characteristics relative to SR are:
 - semantic and physical *interpretation* of major concepts and operations,
 - *reproducibility* of values and results,
 - *availability of measurement* procedures,
 - internal mathematical *consistency* of reasoning,
 - the ways to *evaluate* results.



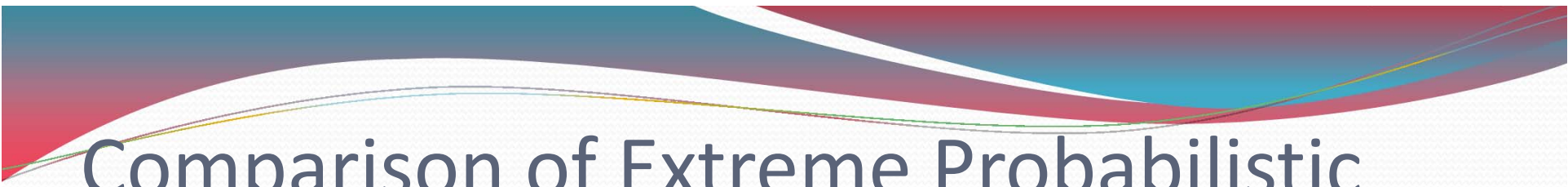
Do we mimic human decision and evaluation processes in fuzzy logic?

- Experimental research did not confirm **initial expectations** about modeling human logic using fuzzy logic based on min and product operations, e.g., [Thole, Zimmerman and Zisno, 1979; Kovalerchuk, Talianski, 1992].
- Therefore, **ad hoc tuning** became common [Kosko, 1997; Nauck et al., 1997; Von Altrock, 1997].
- This made fuzzy logic theory somewhat different from probability theory, where no one tunes in an ad hoc way **definitions of fundamental operations** like union and intersection in the course of a particular study.



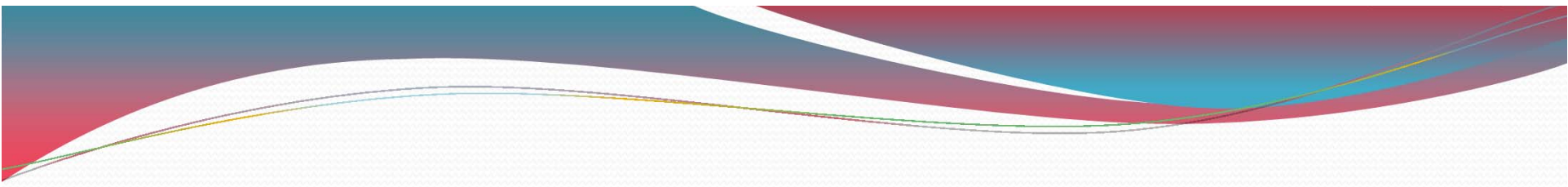
Fuzzy Logic vs. Probability Theory

- Von Altrock, 1997, p.25:
- Especially people working extensively with probability theory have denied the usefulness of fuzzy logic in applications. They claim that all kinds of uncertainty can be expressed with probability theory...
- **Stochastic uncertainty** deals with the uncertainty of whether a certain **event will take place and probability theory lets you model this.**
- In contrast, **lexical uncertainty** deals with **the uncertainty of the definition of the event itself.**
- *“Probability theory cannot be used to model this, as the combination of subjective categories in human decision processes does not follow its axioms”.*
- Stochastic uncertainty and linguistic uncertainty are of different nature.



Comparison of Extreme Probabilistic and Fuzzy Logic positions

<i>Probabilistic Position</i>	<i>Fuzzy Logic Position</i>
All kinds of uncertainty can be expressed with probability theory.	Stochastic and lexical uncertainties have different nature and therefore require different mathematical models.
Probability theory can model stochastic uncertainty, that a certain event will take place.	Probability theory can model only stochastic uncertainty, that a certain event will take place.
Probability theory can model lexical uncertainty with the uncertainty of the definition of the event itself.	Probability theory can not model lexical uncertainty with the uncertainty of the definition of the event itself.
Fuzzy logic may produce useful results, but they currently are based upon weak foundations.	Fuzzy logic often produce useful results
Combination of subjective categories in human decision processes does not follow axioms of fuzzy logic theory.	Combination of subjective categories in human decision processes does not follow axioms of probability theory.
Random sets as part of the probability theory allow modeling all fuzzy logic concepts and reasoning processes.	Fuzzy logic is more general than the probability theory.



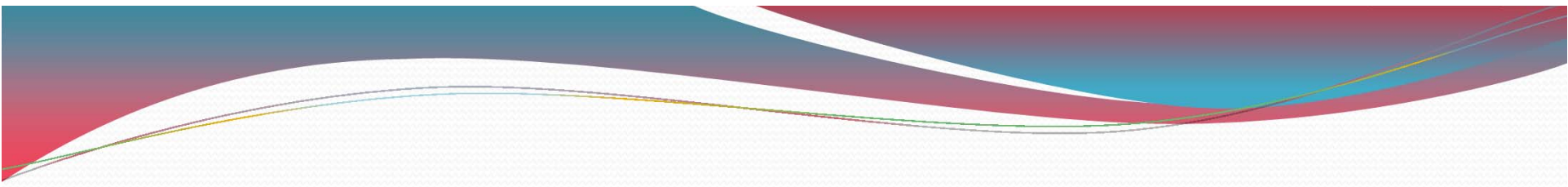
Analysis of extreme statements

- “**Stochastic** and lexical uncertainties have a different nature and, therefore, require different mathematical models.”
- The fact that A and B have different nature does not imply that A and B require different mathematical models.
- Example 1: Water and air have different nature, but both hydrodynamics and aerodynamics are modeled by differential equations that are quite similar.
- Example 2: In ancient time people were able count three stones, three leaves, three trees, that is objects of the same nature. Later people came to an abstract concept “three” for counting objects of different nature.
- The differences in the nature of the entities and stochastic and lexical uncertainties do not mean **automatically** that unified mathematics with them is impossible.
- Stochastic model $\neq$ Probabilistic model.
- Probabilistic models (models that satisfy Kolmogorov's axioms) include stochastic models, subjective probability and random sets based models and many other models.



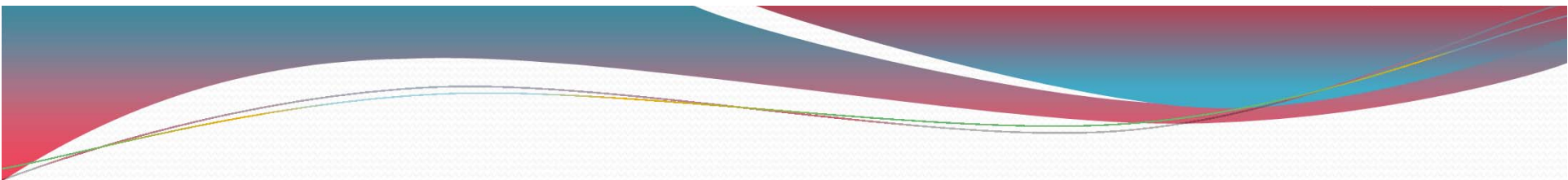
Probability theory and Linguistic uncertainty

- It is true that the probability theory has an origin in stochastic uncertainty as a theory of chances and frequencies back in the 18th century. Nevertheless from 1933 when A. Kolmogorov published an axiomatic probability theory [Kolmogorov, 1956, English reprint], the probability theory moved onto much more abstract level. Actually currently, there are two areas:
- **abstract probability theory** as a part of mathematical measure theory and
- **mathematical statistics** as an area dealing with stochastic uncertainty in the real world.
- These two areas should not be mixed.
- Mathematical statistics matches stochastic uncertainties with abstract probability theory, but it does not prohibit **matching other linguistic and subjective uncertainties with the abstract probability theory**.
- This is actually done with development of **subjective probability theory**, e.g., [Wright, Ayton, 1994].
- Moreover, a **probabilistic linguistic uncertainty theory** was developed [Hisdal, 1998; Kovalerchuk, Shapiro, 1988; Kovalerchuk, 1996a, Kovalerchuk, Vityaev, 2000] inspired by a very productive concept of linguistic variables developed by Zadeh [1977].
- This third approach switches the **focus** from discussing differences in nature of uncertainty to **formalizing contexts of lexical uncertainty**.



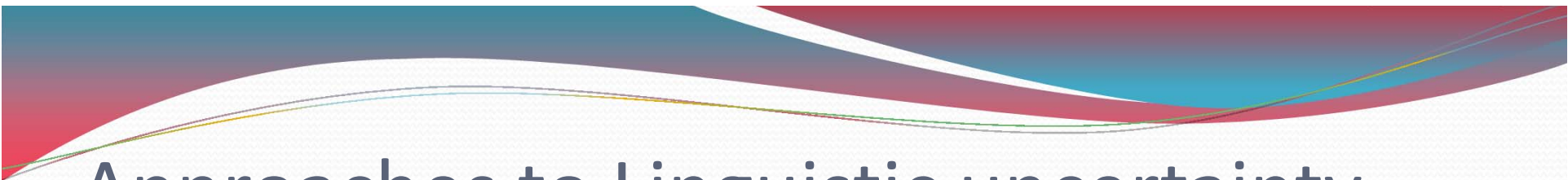
Capturing Linguistic Context

- The formal construction of **probability space** was the principal achievement of probability theory and allowed for the expression of an appropriate **context** for use with probabilities especially as stochastic uncertainties expressed within relative frequencies.
- Clearly, context is important when using linguistic concepts also [Zadeh, 1977] and it is here where fuzzy sets and membership functions typically are used.
- Therefore, we argue that an **analog** (but by no means an identity) of probability space as an expression of context is critically important for firmly founding and more ably advancing the theory and use of fuzzy sets.



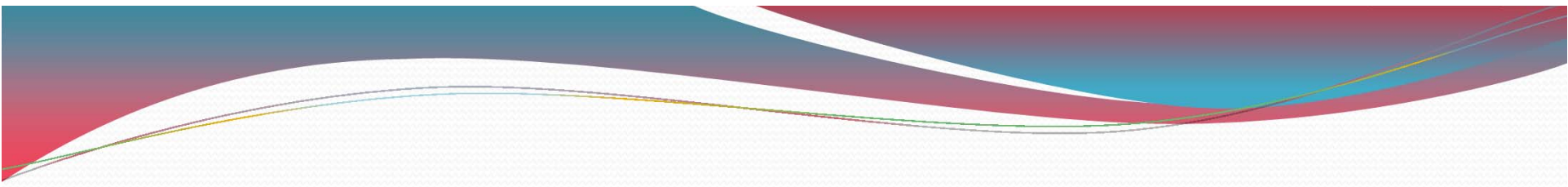
Capturing context

- For at least two centuries, probability calculations were made without a strict context base.
- Many mistakes occurred as a result of such "context-free" calculations of complicated probabilities.
- Currently, the situation for "non-probabilistic inference" (fuzzy logic, et al.) is quite analogous to the one that probability theory was in prior to Kolmogorov's "corrective medicine" which introduced the **concept of probability space**.
- Below we present a way that strengthen foundations of FL with the **concept of linguistic context space**.



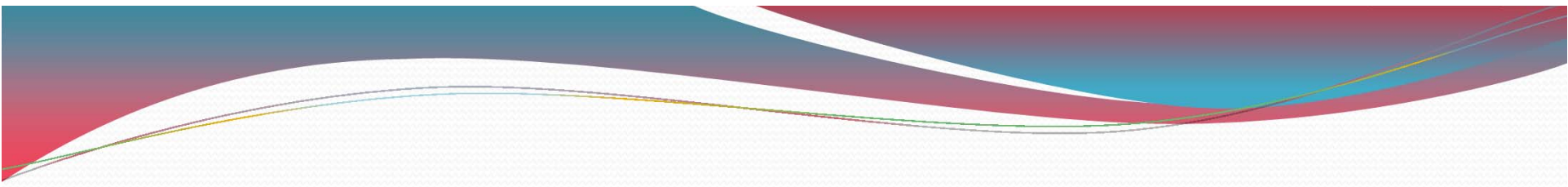
Approaches to Linguistic uncertainty

- L. Zadeh [1988, 2000-2012] supposes that inference under linguistic uncertainty ("lexical imprecision") is precisely the field in which **fuzzy logic** can and should be appropriately and effectively employed.
- P. Cheeseman [1985, 1986] uses **conditional probabilities** and probabilistic measures for specific kinds of propositions.
- Z. Pawlak, et al.[1988] use lower and upper probabilistic approximations as generalizations of **rough sets** to work with linguistic uncertainties.



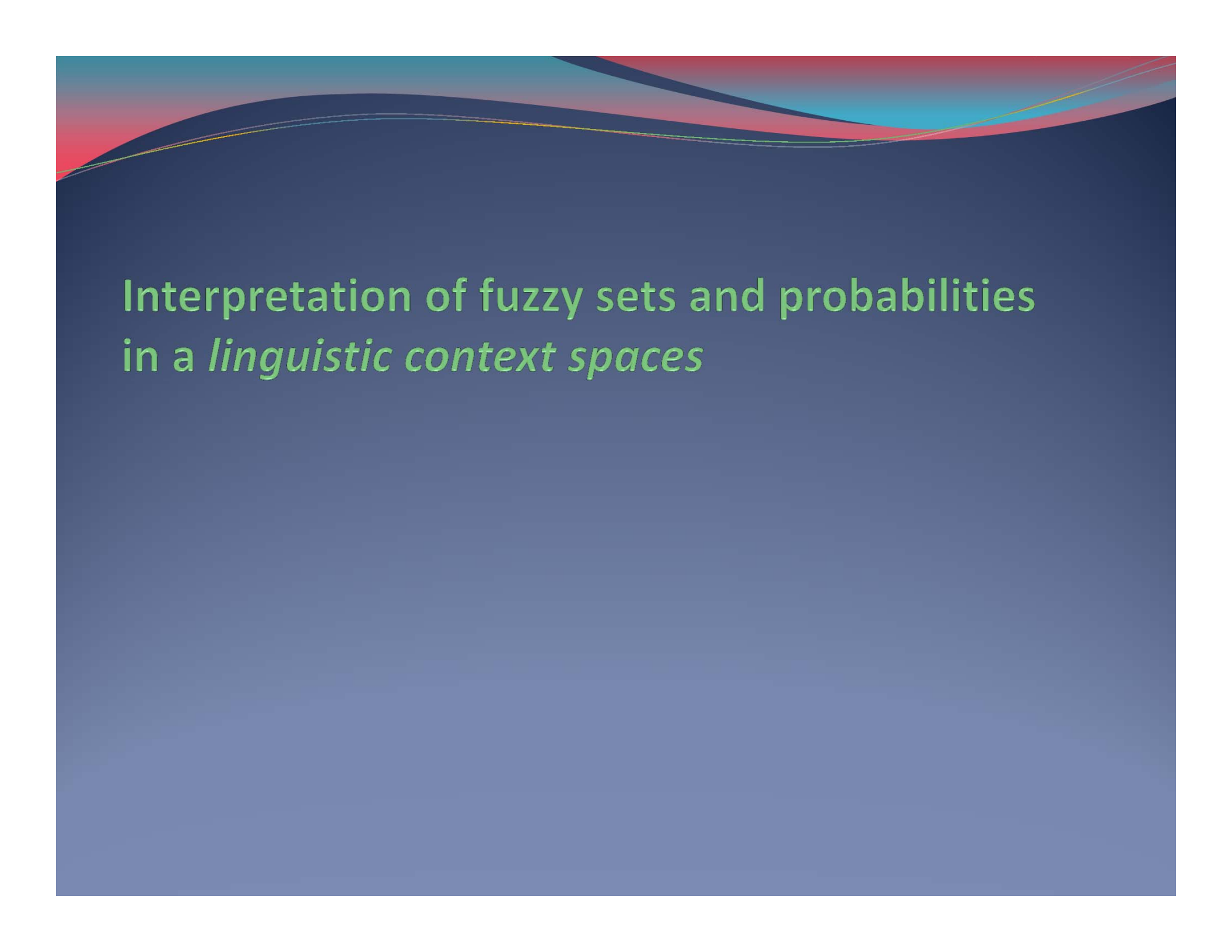
Cheeseman's interpretation

- Cheeseman, 1985:
- The fuzzy logic “possibility distribution” is just the **probability assigned to each of the propositions**
 - “It is possible to put n passengers into Carole's car”
 - for some n , there is considerable uncertainty (i.e., the probability is not close to either 0 or 1).
- "The equivalence of the various fuzzy representations to the corresponding probabilistic representations strongly suggests that the **fuzzy approach is probability in disguise**. However, this is not true because there are major differences in corresponding calculi for computing the values of **logical formulas** from their components.".



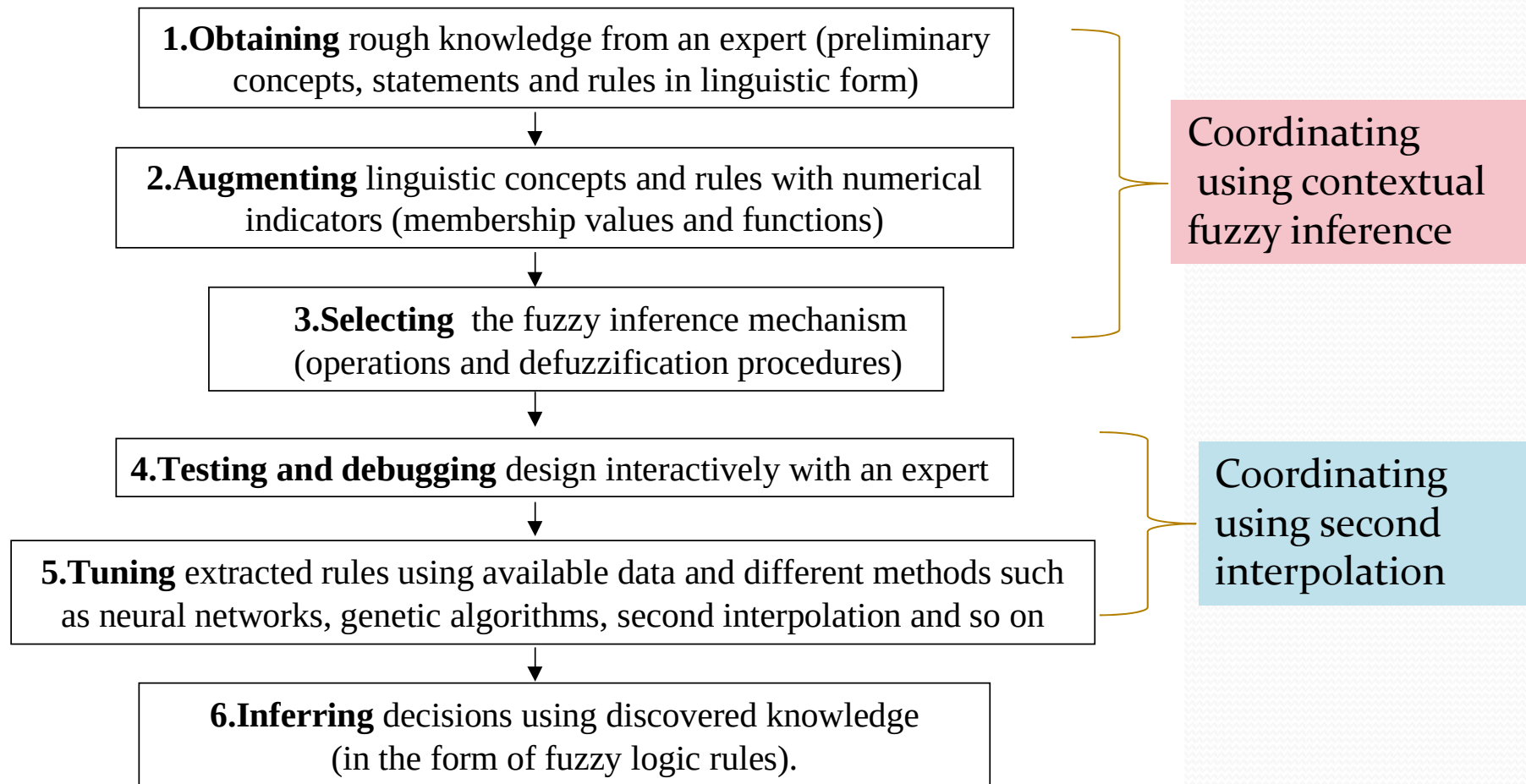
Context space for probability in disguise

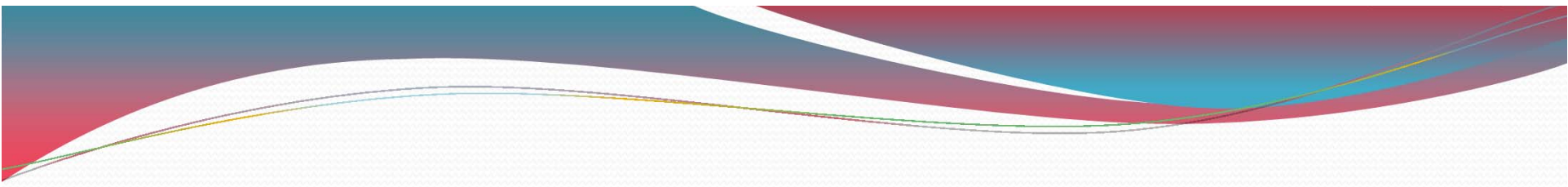
- Cheeseman [1986] did not construct any exact probability space for his example.
 - What is the reference set (set of elementary events) for the needed space?
 - How can we obtain the reference set, probabilities of its elements and their combinations?
 - What are the appropriate distribution and density functions?
- In response to Cheeseman, Dubois & Prade [1990] draw attention to these weaknesses:
 - It is not realistic to find answers for some such questions in many AI applications.
 - Dubois & Prade do not debate Chessman's probabilistic approach, they just assert that
 - In many real cases we do not know **the components of a complete probability space**.
- This may suggest that the fuzzy logic theory addresses this problem as an **uncertainty theory for incomplete probability space**.



Interpretation of fuzzy sets and probabilities in a *linguistic context spaces*

Knowledge discovery and Computing with Words based on Fuzzy Logic





Interpretation vs. accuracy

- Losing meaningful interpretation of rules means losing one of the main advantages of the fuzzy logic approach in comparison with the “black box” data mining and forecasting.
- The **coordinated approach** [Kovalerchuk, Vityaev, 2000] based on concepts of **context space** and **the second interpolation** is a way to **reconcile interpretation and accuracy**.
- Coordinated contextual fuzzy inference using concepts of context space and second interpolation.
 - The first concept addresses coordinating steps 1,2 and 3 and the second one addresses coordinating steps 4 and 5.

Linguistic Context Space for Fuzzy Logic

- Example: Membership functions for interest rate

$$m_{\text{low-rate}}(x), m_{\text{medium-rate}}(x), m_{\text{high-rate}}(x)$$

are not probability density functions with respect to x , but their cross sections for $x=0.04$ (4% rate) could be a probability density function $f_{0.04}(r)$ with respect to r , where r is a linguistic term,

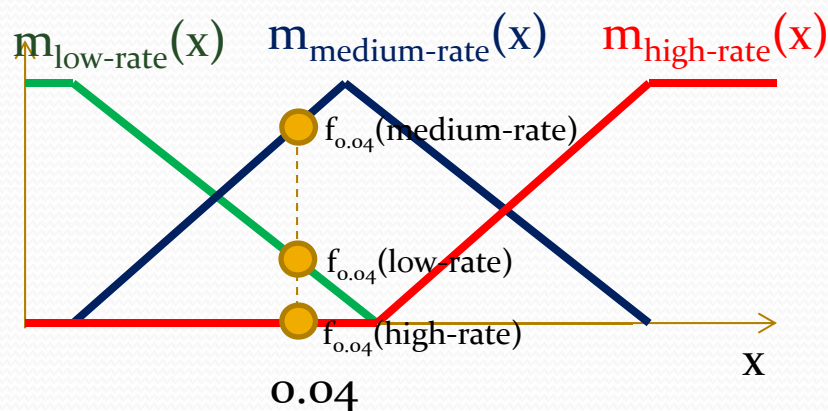
$$r \in \{\text{low-rate, medium-rate, high-rate}\}:$$

$$\text{sum}_{0.04} = f_{0.04}(\text{low-rate}) + f_{0.04}(\text{medium-rate}) + f_{0.04}(\text{high-rate}) = 1,$$

here

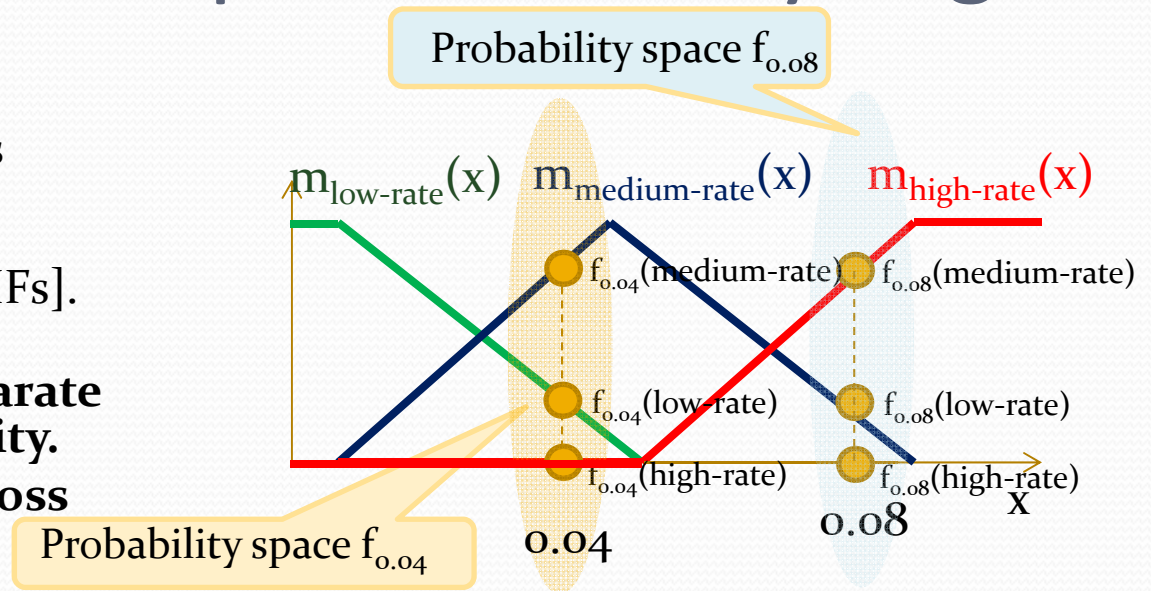
$$\begin{aligned} f_{0.04}(\text{low-rate}) &= m_{\text{low-rate}}(0.04), \\ f_{0.04}(\text{medium-rate}) &= m_{\text{medium-rate}}(0.04) \\ f_{0.04}(\text{high-rate}) &= m_{\text{high-rate}}(0.04). \end{aligned}$$

- We argue that the description of linguistic context, in a natural way, requires more than a **single probability space** alone.
- In this example each interest rate x corresponds to its own small probability space over the *set of linguistic terms* like $\Lambda = \{\text{low-rate, medium-rate, high-rate}\}$.
- [Kovalerchuk, Klir, 1995; Kovalerchuk, 1996; Kovalerchuk, Vityaev, 2000]



Linguistic Context Space for Fuzzy Logic

- Context space is a **class of connected measure spaces represented correctly and compactly by fuzzy logic Membership Functions [MFs]**.
- MF is not a probability distribution, but each separate value of $m(x)$ is a probability.**
- MF is viewed instead as a **"cross section" of probability distributions**.
- Such an interpretation is acceptable for many fuzzy-inference decisions.
- Indeed, experts often construct such **"exact complete context spaces" (ECCSs)** with $\sum_x = 1$ informally for specific problems.



- [Kovalerchuk, Klir, 1995; Kovalerchuk, 1996; Kovalerchuk, Vityaev, 2000]

Context dependence in fuzzy logic and probability theory

- An interpreting mechanism should set **context dependent empirical procedures** for obtaining values of $p(x)$, $m(x)$ and meaningful operations with them.
- Heuristic practice often is “context-free” with little justification, which means that ad hoc tuning is necessary.
- **Context dependence of empirical procedures**
 - includes assigning the value of $m(a)$ in a **co-ordination** with assigning values of m for other entities $b, c, \dots$: $m(b)$, $m(c)$, ...
 - We should not assign $m(\text{tall}, 175 \text{ cm})$ independently of assigning $m(\text{medium}, 175 \text{ cm})$
- **Context independence of computed operations** includes assigning $m(a \& b)$ and $m(a \text{ OR } b)$ truth-functionally:

$$\begin{aligned} m(a \& b) &= f(m(a), m(b)), \\ m(a \vee b) &= g(m(a), m(b)); \end{aligned}$$

i.e., the And and OR operations are functions f and g of only values of $m(a)$, $m(b)$.

Context dependence in fuzzy logic and probability theory

- **The most common operations** in fuzzy logic **min** and **max** representing **AND** and **OR**, respectively, are **context independent**:

$$x \& y = \min(x, y), \quad x \vee y = \max(x, y)$$

where $x = m(a)$, $y = m(b)$.

These operations are also called **truth-functional** [Russell, Norvig, 1995].

- In the context of probability theory:

$$p(a \& b) = p(b)p(a/b),$$

where the conditional probability $p(a/b)$ is a function of the two variables (a,b) , but the $p(b)$ is a function of just one variable.

- **Probability $p(x)$ is context dependent.**
- In $p(a/b)$, b is a context for a ; if b changes, so does the conditional probability. (More in Gaines, 1984; Kovalerchuk, 1990.)

Basic Concepts of Fuzzy Logic

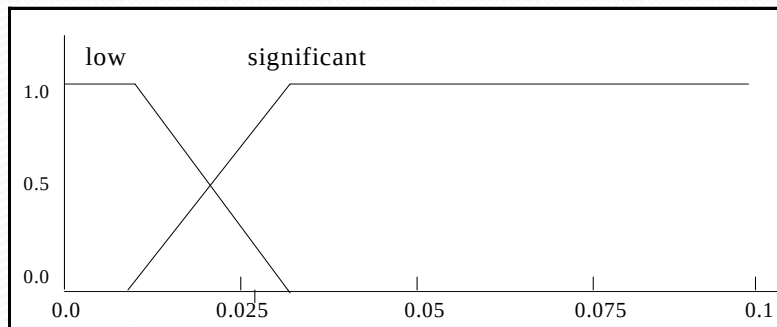
- Example.
 - The number reflecting expert's opinion about **degree of membership** of 0.04 to the set of "low interest rates" is called a value of the **membership function (MF)** $m(x)$ of the fuzzy set "low interest rate".
- Let X be a set of all possible interest rates from 0.0 to 1.0. The set of interest rates $\{x\}$ is called the **universe** X and x is called a "Base variable".
- The **support of the fuzzy set** is a set of x such that $m(x) > 0$.
- The degree of membership m covers the entire interval $[0,1]$, and 0.5 usually corresponds to the most uncertain degree of membership.
- Fuzzy sets generalize conventional "crisp" sets with only two values, 1 and 0. Formally the pair

$$\langle X, \{m_{\text{low-rate}}(x) : x \in X\} \rangle$$

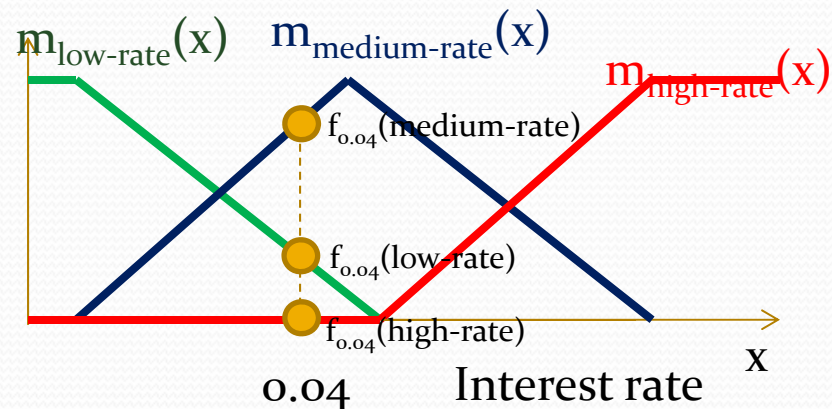
- is called a **fuzzy set of low rates**.
- To define a fuzzy set we need three components: linguistic term ("low rate"), universe ($X=[0,1]$, i.e., all possible rates), and a membership function.
- **Fuzzy logic** combines fuzzy sets **to infer conclusions** from fuzzy logic expressions using membership functions.
- Example: Fuzzy logic assigns a truth value to the composite expression $S_1 \& S_2$
"the interest rate 0.03 is low" AND "the interest rate 0.24 is low"
using $m(S_1) = m(\text{"the interest rate 0.03 is low"}) = 0.72$ and
 $m(S_2) = m(\text{"the interest rate 0.024 is low"}) = 0.82$, $m(S_1 \& S_2) = \min(m(S_1), m(S_2)) = 0.72$

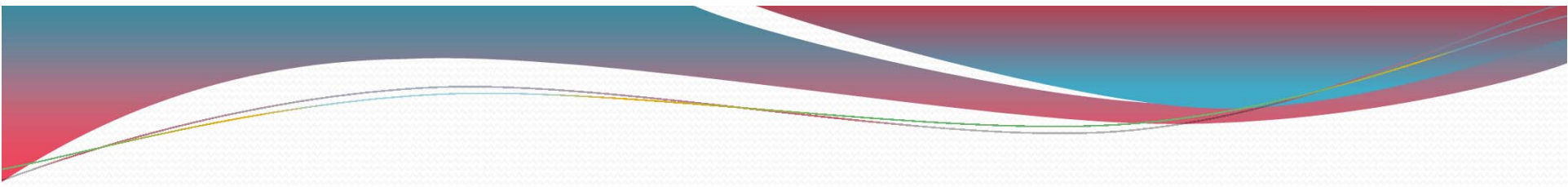
Linguistic variable

- **Linguistic Variables.** The most productive concept of fuzzy logic is the concept of "**linguistic variable**" [Zadeh, 1977].
- Linguistic variable is a set of fuzzy sets defined on the same universe X for related **linguistic terms** like low, medium, high.
- At first glance, any related fuzzy sets can and should be used to create a linguistic variable.
- However, assigning membership degree $m_{\text{significant}}(x)$ for rate x without coordinating this value with already assigned value $m_{\text{low}}(x)$ would be **non-contextual**.



Linguistic variable "trade fee"





Fuzzy Rule Inference

- Rule 1:
IF "interest rate" = low AND "trade fee" = low
THEN
"trade environment" = positive
- Rule 2:
IF "interest rate" = low AND "trade fee" = significant
THEN "trade environment" = positive
- Rule 3:
IF "interest rate" = high AND "trade fee" = significant
THEN "trade environment" = negative
- Rule 4:
IF "interest rate" = medium AND "trade fee" = significant
THEN "trade environment" = indifferent
- The typical fuzzy logic assumption in aggregation is **truth-functionality**: "the truth of complex sentences can be computed from the **truth of the components**."
- Probability combination does not work this way, except under strong **independence** assumptions" [Russell, Norvig, 1995].
- Min and Product are commonly used as AND operation in FL

Inference problems and solutions

- **Rule 1:**

IF "interest rate" = low AND "trade fee" = low THEN "environment" = positive

- **rule 2:**

IF "interest rate" = low AND "trade fee" = significant THEN "environment" = positive

These two rules imply

- **rule 5:**

IF ("interest rate" = low OR "interest rate" = medium) AND "trade fee" = low
THEN "environment" = positive

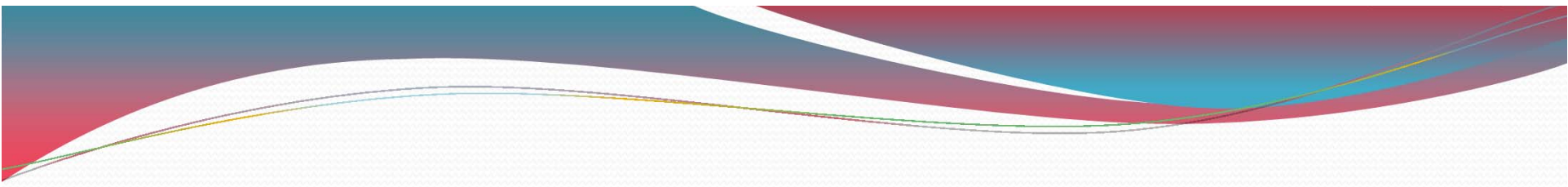
- Notation: R is interest rate R and F is trade fee.

- Example:

- $R=0.04$, $F=0.01$, $m_{\text{low}}(R)=m_{\text{medium}}(R)=0.5$ and $m_{\text{low}}(T)=1$, i.e., F is a low fee and R as between low and medium interest rates.

- According to rule 5 the **environment** for this pair $\langle R, T \rangle$ is **positive**, i.e., $m_{\text{positive-environment}}$ should be 1.

- However, traditional fuzzy MIN-MAX inference used in [Von Altrock, 1997]) produces $m_{\text{positive-environment}}(R, F) = 0.5$ that is the highest level of uncertainty of the conclusion.



Fuzzy Rule debugging

- An alternative Bounded sum (BSUM) defuzzification available in Fuzzytech has produced
 - $MF_{\text{positive-environment}}(R.T)=1$.
- This output is fully consistent with logical and intuitive expectations.
- While wrong results can be fixed error-prone debugging of very heuristic procedures, a thorough **analysis of context space** is a better way to get a correct result.
- This can be done by using the **second interpolation** approach of fuzzy set theory (Kovalerchuk [1996, 2000]).

Rule debugging

- Even if debugging on the intermediate step has shown consistency of the rules they still may be inconsistent in a further inference.

- Example:

- **Rule 2:**

IF "interest rate" = low AND "trade fee" = significant THEN "environment" = positive

- **Rule 4:**

IF "interest rate" = medium AND "trade fee" = significant THEN "environment" = indifferent

- Logically these rules produce

- **Rule 6:**

IF ("interest rate" = low OR "interest rate" = medium) AND "trade fee" = significant THEN "environment" = indifferent

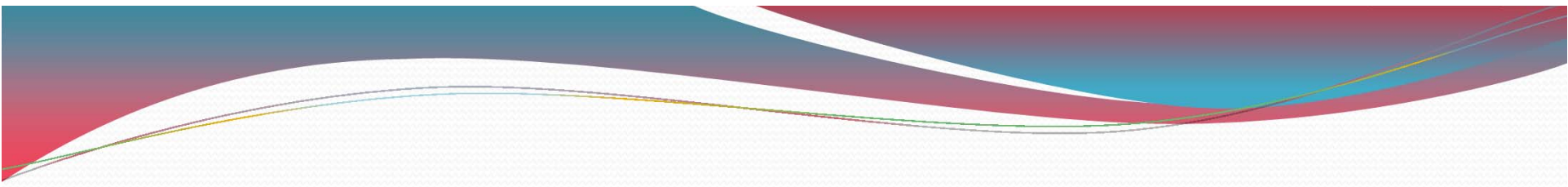
- For interest rate $R=0.04$ and fee $F=0.03$, i.e., exactly between picks of membership functions for low and medium interest rates,

$$m_{\text{low-interest}}(R) = m_{\text{medium-interest}}(R) = 0.5, \quad m_{\text{significant-fee}}(F) = 1.$$

F definitely represents a significant fee and R is something between low

and medium interest rate.

- FuzzyTech with MAX-MIN result defuzzification delivered intuitively and logically consistent results for positive and indifferent trade environments separately



Rule debugging

- Joint consideration of positive and indifferent trade environments produces inconsistent results.
- Rule 6 is an OR combination of positive and indifferent trade environments.
- According to rule 6 $MF_{\text{positive-or-indifferent-environment}}(R,F)$ should be 1, but MAX-MIN fuzzy inference can produce only 0.5.
- Another inference, which uses BSUM handles this situation correctly. Nevertheless, it does not mean that BSUM is the right choice.
- An adequate analysis should involve a careful study of context [Kovalerchuk, 1996]. One of the reasonable **alternatives to heuristic debugging** is the use of **second interpolation based on picks of membership functions** and values of output variable for these picks.

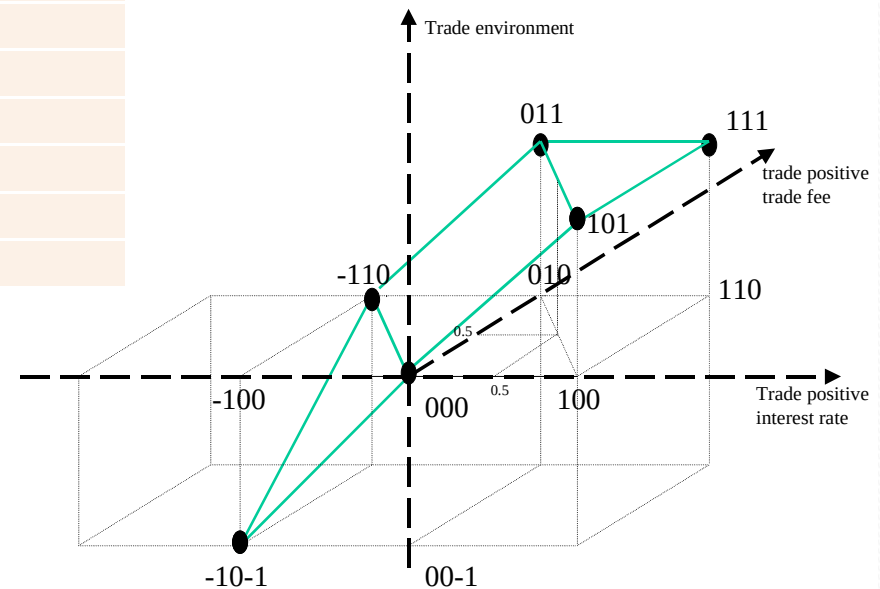
Second interpolation

If		Then
Interest Rate	Trade Fee	Environment
Low	Low	Positive
Medium	Low	Positive
High	Low	Indifferent
Low	Significant	Positive
Medium	Significant	Indifferent
High	Significant	Negative

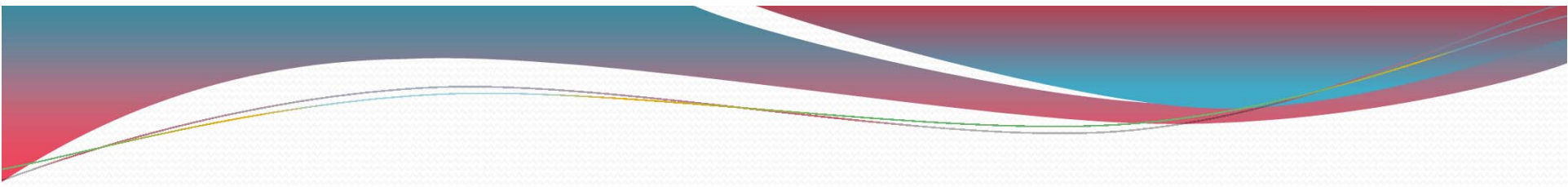
Matrix of linguistic rules

If		Then
Interest Rate	Trade Fee	Environment
1	1	1
-1	0	-1
0	0	0
1	0	1
0	1	1
-1	1	0

Numerical rule table



Surface presentation of rules

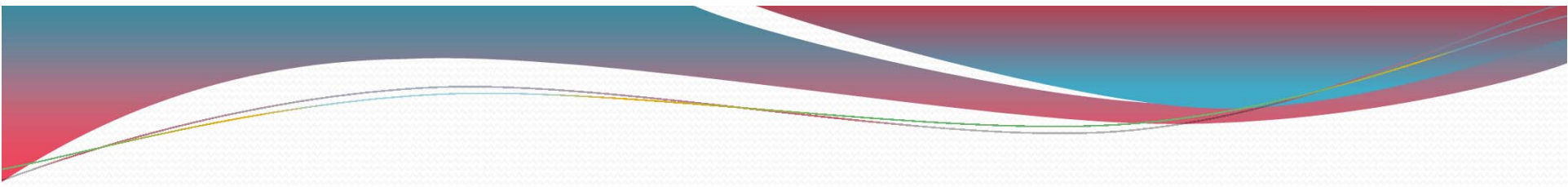


Second interpolation

- Six rules presented are converted into points in three-dimensional space.
- MF values for interest rate, trade fee and environment all equal to 1 and that if the first two MFs are equal to 1 then the third should be 1 also.
- Use axes for the interest rate, trade fee and trade environment.
- This also shows the simple surface connecting these points representing the degree of positive environment for every value of the interest rate and trade fees not only integer values. This means a defuzzified output is delivered directly.
- There are simple formulas to define this surface analytically. The given surface does not suffer from problems of inconsistency known for standard fuzzy inference.
- The surface delivers $MF_{\text{positive-environment}}(R, F) = 1$, for $MF_{\text{low-interest}}(R) = 0.5$ and $MF_{\text{low fee}}(F) = 0.5$, see Figure 7.10.
- [Kovalerchuk, 1996b]

Constructing coordinated contextual fuzzy sets and linguistic variables

- **Example 1 (Inconsistency).**
- Let the *first step* (obtaining a linguistic statement) be ended with
 - **statement S1:** “A person of age 59 is Almost Old OR Old”.
- Then in the *second step* the linguistic concepts “Almost Old” and “Old” are augmented with numerical indicators of their truth for age 59:
 - $m(59 \text{ is Old}) = 0.55$,
 - $m(59 \text{ is Almost Old}) = 0.45$,
 - $m(59 \text{ is Old OR } 59 \text{ is Almost Old}) = 1.00 \text{ (true)} \quad (1)$
- Assume also that the *third step* ends with the standard fuzzy logic max operation for OR. it produces:
 - $m(59 \text{ is Old OR } 59 \text{ is Almost Old}) = \max\{OL(59), AO(59)\} = \max\{0.55; 0.45\} = 0.55, \quad (2)$



Example of inconsistency

- Truth of the statement S_1 ="59 is Old OR 59 is Almost Old" is very questionable (truth degree 0.55). Nevertheless, more naturally this statement should be true or nearly true.
- Therefore, the *fourth step* (debugging) is needed, e.g., substituting max for SUM.
- Here MF values are supported intuitively and experimentally [Hall, et al, 1986], reflecting the given linguistic statement: "A person of age 59 is Almost Old or Old", but the choice of the max operation is not.
- Unfortunately, this is a standard practice in fuzzy inference. In other words, independently defining a membership function m and OR operation as max created inconsistency.
- The negative aspect of debugging is that it is done ad hoc and no one can guarantee that with more age concepts such as "Not Old" or "Very Old" a similar problem will not raise.
- The concern is that SUM could be just a first step in debugging and we can not universally substitute max for SUM. This is confirmed in the next example.

Constructing coordinated contextual fuzzy sets and linguistic variables

- **Example 2 (Case sensitive debugging).**
- Assume that we have substituted max operation for sum operation in the debugging of Example 1. Consider **statement S2**:
 “A person of age 59 is Almost Old OR Old OR Very Old OR Not Old”
 as a true statement (Step 1).
- Step 2 produced according to [Hall, et al., 1986]:
 - $m(59 \text{ is Old}) = 0.55$,
 - $m(59 \text{ is Almost Old}) = 0.45$,
 - $m(59 \text{ is Not Old}) = 0.35$,
 - $m(59 \text{ is Very Old}) = 0.06$.
- Step 3 delivers for statement 2:
 $OL(59) + AO(59) + NO(59) + VO(59) = 0.45 + 0.35 + 0.55 + 0.06 = 1.41 \quad (3)$
- Thus, the truth-value 1.41 in (3) instead of 1.0 also should be debugged.

Constructing coordinated contextual fuzzy sets and linguistic variables

- **Example 3 (Case sensitive debugging).**

- Step 1. Consider

IF the age of 32 years is Not Old (NO) THEN the age of 59 is either Almost Old OR Old.

Intuitively, this statement seems to be approximately true.

- Step 2. Hall et al [1986] provide the following experimental values:

$$NO(32)=0.95$$

$$AO(32)=0.04$$

$$OL(32) = 0.01.$$

Also we assume from (1) that $m(\text{age of 59 years is AO OR OL})=1$.

- Step 3. Use logical equivalence $(C \Rightarrow D)$ and $(\text{not } C) \vee D$.

where $D = \{\text{age of 59 years is AO OR OL}\};$

$(\text{not } C) = \{\text{age of 32 years is AO OR OL}\}.$

- This negation **differs** from a simple negation of $\{\text{age of 32 years is NO}\}$, i.e., it is not simply $\{\text{age of 32 years is OL}\}$. For the other context of $(\text{not } C)$, the other set of linguistic terms may take place.

Constructing coordinated contextual fuzzy sets and linguistic variables

- Using the standard max-min fuzzy logic approach (max for OR and min for AND):

$$m(E) = m((\text{not } C) \vee D)$$

$$= m(\{\text{age of 32 years is AO OR OL}\} \vee \{\text{age of 59 years is OL}\})$$

$$= \max(m(\{\text{age of 32 years is AO OR OL}\}), m(\{\text{age of 59 years is AO OR OL}\}))$$

$$= \max(\max[m(\text{age of 32 years is AO}), m(\text{age of 32 years is OL})],$$

$$\max[m(\text{age of 59 years is AO}), m(\text{age of 59 years is OL})])$$

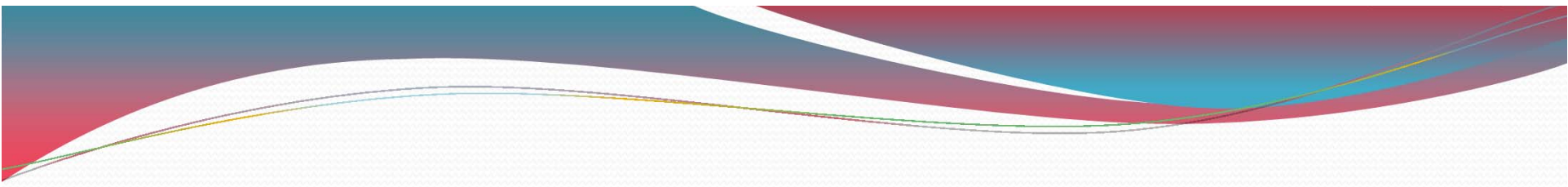
$$= \max(\max(0.04, 0.01), \max(0.55, 0.45)) = 0.55.$$

- Thus, the min-max operation gives an answer ($m=0.55$) that certainly is NOT intuitively correct.
- The reason for inconsistency is the same -- max-min operations were selected in step 3 without coordinating with steps 1 and 2. In addition, the use of the sum produces results, which should be debugged too ($m>1$).
- Examples 1-3 illustrate the coordination problem of linguistic statements and rules.
- It is also important to note that **probability theory does not require a debugging** operation in similar situations. This is illustrated in Example 4 below.

Example 4:

Linguistic fuzzy-probabilistic approach

- Task: Estimate
 $m(\text{age of 32 is NO AND age of 59 is AO})$.
- Consider the Cartesian product of the two probability spaces:
 $\{\text{NO, AO, OL}\} \times \{\text{AO, OL}\}$ that consists of 6 pairs
 $\{(\text{NO,AO}), (\text{AO,AO}), (\text{OL,AO}), (\text{NO,OL}), (\text{AO,OL}), (\text{OL,OL})\}$.
- Compute MFs of those pairs with the first component for age 32 years and the second component for age 59 years.
 $\text{NO}(32) = 0.95$, and $\text{AO}(59) = 0.55$ [Hall, et al., 1986].
Under common independence assumption Cheeseman [1985, 1986] :
 $m(32 \text{ is NO AND } 59 \text{ is AO}) = \text{NO}(32) * \text{AO}(59) = 0.95 * 0.55 = 0.52$.
- Compute $m(E) = m((\text{not C}) \vee D)$
 $= m(\{\text{age of 32 is AO OR OL}\} \vee \{\text{age of 59 years is AO OR OL}\}) =$
 $= [\text{AO}(32) + \text{OL}(32)] + [\text{AO}(59) + \text{OL}(59)] - [\text{AO}(32) + \text{OL}(32)] * [\text{AO}(59) + \text{OL}(59)]$
 $= [0.04 + 0.01 + 0.55 + 0.45] - [0.05 * 1.00] = [1.05] - [0.05] = 1.00$.
- This **intuitively acceptable output** ($m(E)=1$) **does not requires debugging** of standard probabilistic operations for independent events – “+” for OR and “*” for AND. If the independence assumption is not true, then we should use conditional probability:
 $\text{NO}(32) * \text{AO}(59 / [\text{NO}(32)]) = 1.00$.
Here,
 $\text{AO}(59 / ([\text{NO}(32)]) = 1.00$
means that 59 years of age IS "Almost Old", GIVEN THAT 32 years of age IS "Not Old".



Example 5 (Dependent terms)

- Compute truth-value for the statement:

B = (age of 32 years is NO *AND* age of 59 years is AO)

OR (age of 32 years is NO *AND* age of 59 years is NO)

OR (age of 32 years is NO *AND* age of 59 years is OL).

- Use of standard probabilistic operations under the same supposition of independence produces:

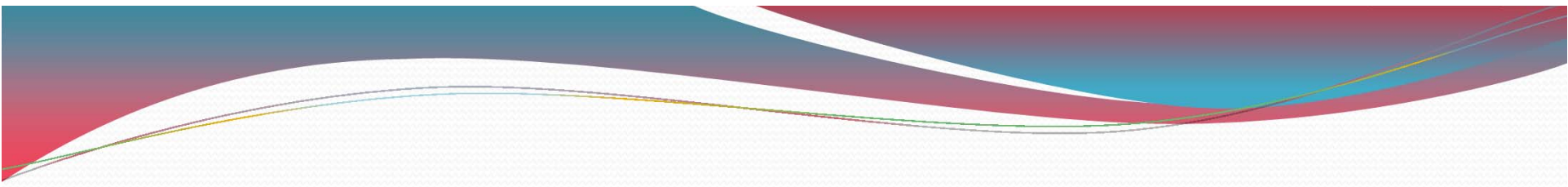
$$m(B) = (0.95 \cdot 0.55) + (0.95 \cdot 0.35) + (0.95 \cdot 0.45) = 0.52 + 0.33 + 0.42 = 1.27,$$

i.e., debugging is needed to get a truth value not greater than 1.

- Conclusion: a **simple substitution of fuzzy operations for probabilistic operations does not solve the problem of debugging completely.**
- It must be **context dependent**. The term Not Old can not be considered as independent from terms Almost Old and Old for age 59 as shown in experiments [Hall et al, 1986].

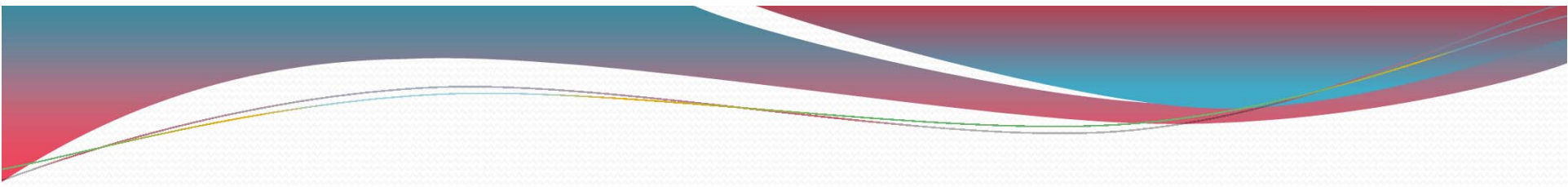
Example 6 (nested, complementary and overlapping linguistic terms)

- **Step 1.** Introduce two sets of terms for the ages of 32 years and 59 years:
 $T(32)=\{NO, AL\}$, $T'(32)=\{NO, NO \text{ OR } AL\}$ and $T(59)=\{AO, OL\}$, $T'(59)=\{AO, AO \text{ OR } OL\}$
Assume that two statements are true
Age of 32 is (Not Old OR (Not Old OR Almost Old)),
Age of 59 is (Almost Old OR (Almost Old OR Old))
For simplicity, NO OR AL will be denoted as NOAL and AO OR OL will be denoted as AOOL.
- **Step 2.** Let $NO(32)=0.95$ and $AO(59)=0.55$ as in previous examples. It is also assumed that
 $m(\text{age 32 is NOAL})=1$,
 $m(\text{age 59 is AOOL})=1$.
- We will call $T(32)$ and $T(59)$ **exact complete term sets** for 32 and 59, respectively, because of these properties. The $T(32)$ and $T(59)$ terms express **distinct concepts** and the prime terms in $T'(32)$ and $T'(59)$ express **nested concepts**. The set of exact complete terms is the base concept of an **exact complete context space**.
- **Step 3.** Select sum to represent OR operation.
The sums for nested T' are more than 1:
 $m(\text{Age of 32 is Not Old OR (Not Old OR Almost Old)})=0.95+1=1.95$
 $m(\text{Age of 59 is Almost Old OR (Almost Old OR Old)})=0.55+1=1.55$.
Debugging is needed for T' to suppress these sums to 1.
- **Step 4.** One way of debugging is to return to max operation:
 $m(\text{Age of 32 is Not Old OR (Not Old OR Almost Old)})=\max(m(\text{age of 32 years is NO}), m(\text{age of 32 years is NO OR AL}))$
 $=\max(0.95, 1)=1$.
Similarly,
 $m(\text{Age of 59 is Almost Old OR (Almost Old OR Old)})=\max(m(\text{age of 59 years is AO}), m(\text{age of 59 years is AO OR OL}))$
 $=\max(0.55, 1)=1$.



Analysis of Example 5

- Now we have obtained a corrupt loop -- to debug max we moved to the sum and to debug the sum we moved to max again.
- Thus, simple “**context-free**” **debugging** trying different operations helps to solve one problem, but can create another.
- **Contextual debugging** or better **contextual designing** steps 1-4 is needed. “Contextual debugging here would begin with noticing that incorrectness is a result of “**context free**” computations.
- **Distinct and nested terms** have been mixed in a single context space and the same space and term set was used for different ages (32 and 59 years).
- Nested terms may require min operation, but distinct terms may require sum.
- In the statement “A person of age 59 is Almost Old OR Old OR Very Old OR Not Old” terms Almost Old and Old are **distinct** ones, but terms Old and Very Old are **nested**. For **nested** terms, if someone is called Very Old he/she also can be called Old, i.e., here Very Old is interpreted as nested to Old. This is not the case for **distinct** terms Almost Old and Old. We can not say that if someone is Almost Old he/she is also Old or if he/she is old that he/she is almost old.



Analysis of Example 5

- In the mixed space with nested and distinct terms, **there is no way to find one “context-free” operation in the process of debugging.** Switching between these operations would be needed.
- However, this is only a partial solution. The number of operations and switches can be as large as the number of statements in the domain.
- To avoid all these complex problems it would be more natural to **develop a distinct or a nested context space in advance.**
- Next, we discuss **the** methods of developing these spaces.

Analysis of Example 5

- Consider a statement “A person of age 59 is Almost Old OR Old OR Very Old OR Not Old” and associated nested terms
“Almost Old OR Old OR very Old OR Not Old” $\supset$ “Almost Old OR Old OR very Old” “Old OR very Old” $\supset$ “Old” $\supset$ “Very Old”.
- If m given experimentally only for “Old” and “Very Old” then we cannot compute m for “Almost Old OR Old OR very Old OR Not Old” without additional information.
- Getting additional information. Transform the term set from the statement “A person of age 59 is Almost Old OR Old OR Very Old OR Not Old” in to the set of **distinct terms in "exact complete context space (ECCS)"**.

Consider terms C_1 - C_4 :

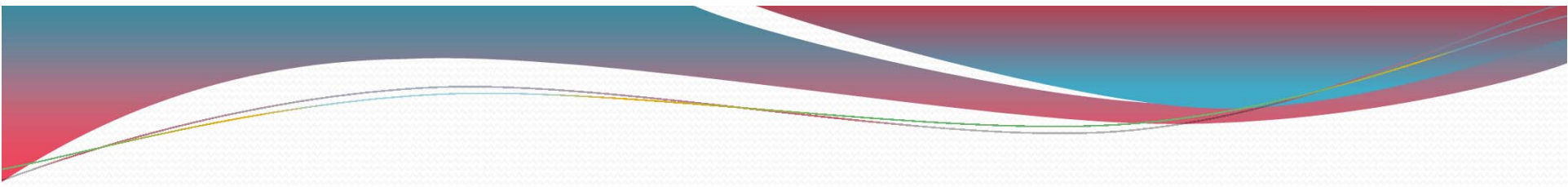
C_1 ="Almost Old",

C_2 ="Old",

C_3 = "Very Old and not Old" and

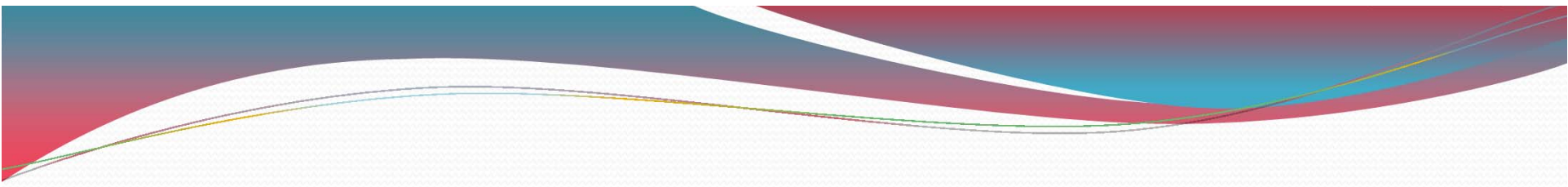
C_4 ="Not Old" and not ("Almost Old OR Old OR Very Old").

- C_1 and C_2 form a complete space for age 59 [Kovalerchuk, Vityaev, 2000]
- This imply:
$$\begin{aligned} m(\text{"A person of age 59 is Almost Old OR Old OR Very Old OR Not Old"}) \\ = m(\text{"A person of age 59 is Almost Old"}) + m(\text{"A person of age 59 is Old"}) \\ + m(\text{"A person of age 59 is "Very Old and not Old"}) \\ + m(\text{"A person of age 59 is Not Old and not(Almost Old OR Old OR Very Old)}) \\ = 0.55 + 0.45 + 0 + 0 = 1. \end{aligned}$$



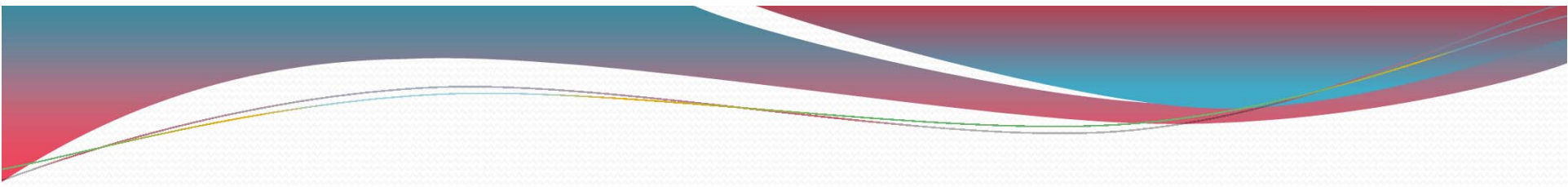
Context space

- Examples above show that the max-min fuzzy logic approach produces the same result as the "exact complete" context space approach if (and only if) nested spaces are accurately constructed with a one-to-one correspondence to the ECCS.
- Thus, whenever natural language terms are mostly nested for a given applied problem, we should try to construct the complete nested term set (sets) within the whole context. That is, the entire context should be equivalent to an ECCS. The correspondence between ECCS and a **nested context** is close to correspondence between a **density function** and a **distribution function** in probability theory.
- The above examples should suffice to show that probability spaces could be constructed for the linguistic uncertainties usually given empirically in situations involving fuzzy sets. Thus, in this regard, Cheeseman's probabilistic position [Cheeseman, 1985, 1986] is justified. We should also emphasize that this is a typical situation for many applications of fuzzy logic. We can construct exact complete context space(s) (as in examples, above) without serious problems, and such contexts likely do not change during their "lifetimes".
- On the other hand, as Dubois and Prade noted [1990], in dealing with knowledge base problems, it is not uncommon to have to deal with the complete probability spaces. Thus, in such cases, it is **often impossible to specify probability spaces**, i.e., they may well be dynamically changing and/or do not provide a "complete" context. However, it will be important to have some real examples from knowledge base applications where it is truly impossible to construct "complete" context spaces. Such examples would provide a justification of the Dubois and Prade "impossibility" position [1990].



Context space

- Nonetheless, the examples above show that without specification of "exact" complete context space, many mistakes can be made in calculating compound membership functions.
- The complete context is of importance and has considerable applied interest for fuzzy inference.
- The "over completeness" of natural language can (and should be) "corrected" ("debugged") for the examples with $\text{sum} > 1$.
- Such a correction is especially effective whenever we construct artificial linguistic variables in fuzzy inference applications.
- An "**over complete**" space will be generated if we will use a linguistic space {NO, AO, OL} for both ages and all pairs:
$$\{\text{NO}, \text{AO}, \text{OL}\}_{32} \times \{\text{NO}, \text{AO}, \text{OL}\}_{59}$$
- The term Not Old (NO) is redundant for age 59 and term Old (OL) is redundant for age 32.
- For "incomplete" spaces, we
 - cannot calculate m for complex expressions because not all components are available;
 - should not use independence because we "don't know" what else to assume, and
 - cannot give equal probabilities to additional formal elements of the space, as was shown by Dubois and Prade.



Context space

- Define **context space**.

- Example:

$$OL(59) + AO(59) = 0.55 + 0.45 = 1.00 \text{ [Hall, et al. 1986]}:$$

where: $OL(59) = m(59/\text{Old})$, and $AO(59) = m(59/\text{Almost Old})$.

- The pair {AO, OL} provides a "**complete**" ("**exact complete**") **context** for the age of 59 years.
- Let us add a new term Not Old (NO). {NO, AO, OL} would be designated as an "**over complete**" **context**, and the context designated by {OL} would be an "**incomplete**" **context**.
- For the age of 32 years:

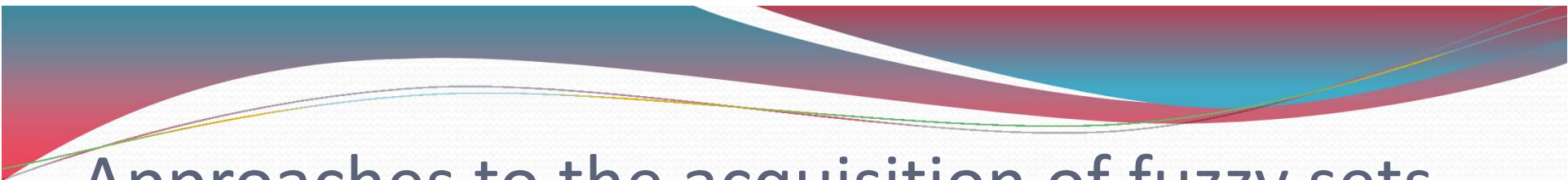
$$NO(32) + AO(32) + OL(32) = 0.95 + 0.04 + 0.01 = 1.00$$

or

$$NO(32) + AO(32) + OL(32) + VO(32) = 1.00,$$

where VO stands for Very Old.

- The contexts {NO, AO, OL} and {NO, AO, OL, VO} are "exact complete context spaces" (ECCS) for the age of 32 years.
- Based on Hall, et al. [1986], contexts are incomplete for ages 18 and 24 years, and are over complete for ages 40, 51, 59, 68, 77, e.g.,
 $OL(59) + NO(59) + AO(59) + VO(59) = 0.45 + 0.35 + 0.55 + 0.06 = 1.41$.
- More formal definition of a context space [Kovalerchuk, 1996].



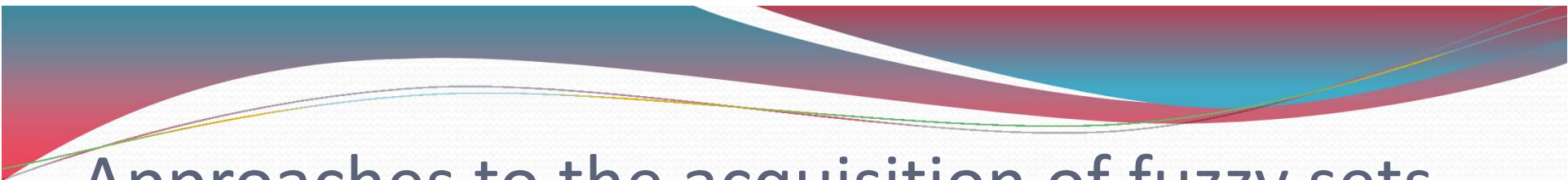
Approaches to the acquisition of fuzzy sets and membership function

- One of the most important problems in fuzzy sets theory is obtaining appropriate MFs for a given **context space**.
- Semantic operational procedures [Hisdal, 1998] and modal logic [Resconi et al, 1992,1993] are preferable to other available techniques.
- When preferable operational ways are used jointly with the concept of ECCSs then meaningful fuzzy logic conclusions are produced resolving problems illustrated in the examples above.
- Is it true that fuzzy sets can be constructed easier than probability distribution? Fuzzy logic literature often claims simplicity as its advantage.
- We show that this is true only if we ignore contextual relations (context space) and pay for that with the complex and time consuming debugging of fuzzy sets later.



Approaches to the acquisition of fuzzy sets and membership function

- “Fuzzy set theory does not require the sum of m values on X to be equal to 1.00”. Indeed, for many purposes there is no problem if this sum exceeds 1.00. Clearly, however, the important question is what we want to do with m values.
- Example: let $m_A(x) = 0.70$, and $m_A(y) = 0.90$; if we only want to conclude that the “degree of belief” for x is less than that for y for A to be the case, then there is no problem with the sum here being > 1.00 .
- Here any monotone transformation of the initial m values will preserve the preference order, e.g., $m_A(x) = 0.95 < m_A(y) = 1.20$.

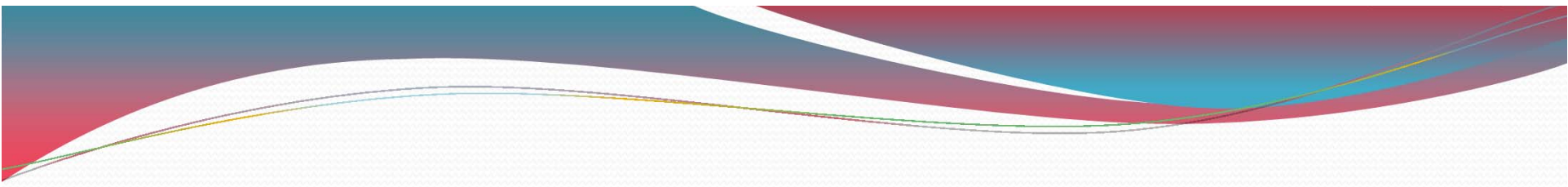


Approaches to the acquisition of fuzzy sets and membership function

- Fuzzy statistics [Hall, *et al.*, 1986; Hisdal, 1984] provide several approaches to the acquisition of MFs. Let's consider the following:
- a) Values for specific linguistic statements such as "x is a low rate" are obtained by asking members of a group to register their agreement/disagreement with the statement. Responses are **binary** (i.e., "yes" or "no").
- b) Subjects are asked to rate their agreement/disagreement with the statement on a **continuous scale** of 0 to 1. Then, using mathematical techniques, statistical characteristics of the group of subjects are computed. Values obtained in this way are sometimes referred to as "**fuzzy expected (average) values**" (FEVs).
- Hisdal [1998] introduced **semantic operational procedures** for defining fuzzy sets. It is necessary here for at least two persons to participate: The "**knowledge engineer**" (E), who gives instructions to the **subject(s)** S ($S_1, S_2, \dots$), and every procedure is performed on a set of **objects**. For example, the "objects" could also be people. Every semantic procedure must have a **reference label set** $\Lambda = \{\lambda\}$. For example, $\Lambda = \{\text{young, middle-aged, old}\}$, and the λ_i 's are called labels.
- **Labeling (LB) procedure.** In psychophysics, this is a forced- (multi-) choice procedure.
- Example: E instructs S to answer the question "**What is today's interest rate?**" by choosing one of the labels -- $\Lambda = \{\text{low, medium, high}\}$. (The requirement of referring to Λ pertains to the next three (3) definitions as well.)

Approaches to the acquisition of fuzzy sets and membership function

- Example: E instructs S to answer the question "What is John's age?" by choosing one of the labels -- $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_i, \dots, \lambda_L\}$. (The requirement of **referring to Λ** pertains to the next three definitions as well.)
- **Yes-No (YN) procedure.** In psychophysics, this is **forced two-choice procedure**.
- Example: E asks, "Is today's rate high?" S answers either "yes" or "no".
- Example: E asks, "Is John old?" S answers either "yes" or "no".
- **LB-MU procedure.** In psychophysics, this is a kind of **direct scaling procedure**.
- Example: E asks S to **specify the "degree" (L) to which each of the labels from Λ is appropriate for today's interest rate** (i.e., i.e., $m_{\lambda_i}(\text{today}) \in [0,1]$, where
- $m_{\lambda_1}(\text{today}) = m_{\text{low-int-rate}}(\text{today})$, $m_{\lambda_2}(\text{today}) = m_{\text{medium-int-rate}}(\text{today})$ and
- $m_{\lambda_3}(\text{today}) = m_{\text{high-int-rate}}(\text{today})$).
- Example: E asks S to specify the "degree" (L) to which each of the labels from Λ is appropriate to John (i.e., $m_{\lambda_i}(\text{John}) \in [0,1]$).
- **YN-MU procedure.** In psychophysics, this is a kind of **direct scaling procedure**.
- Example: E instructs S to **specify the degree $m_{\text{yes-high-rate}}(\text{today}) \in [0,1]$ to which the answer "yes" is correct to the question "Is interest rate high today?"**
- Example: E instructs S to specify the degree $m_{\text{yes-young}}(\text{John}) \in [0,1]$ to which the answer "yes" is correct to the question "Is John young?"



Obtaining linguistic variables

The process of obtaining linguistic variables should consist of the following steps:

- Step 1. Determine a reference set Λ .
- Step 2. Define initial type of the reference set Λ -- distinct, nested or mixed.
- Step 3. Correct Λ to bring it into distinct or nested type.
- Step 4. Obtain MFs for all terms of corrected Λ using semantic or modal logic operational procedures.
- Step 5. Define the type of the corrected reference set Λ , i.e., distinct, nested or mixed.
- Step 6. Correct Λ again to bring it into distinct or nested types repeating steps 4,5 and 6 until getting an acceptable exact complete context space.

“Debugging” and tuning MFs and operations

- Often selection, debugging and tuning are done by "blind" optimization.
- Simultaneous optimization of both Input/Output interfaces and a linguistic subsystem without integrity constraints can generate meaningless linguistic terms [Pedrycz, Valente de Oliveira, 1993].
- **How to eliminate heuristic procedures in debugging and tuning?**
- **How to simplify the related computations?**
- **Heuristic debugging** is trying different fuzzy operations.
 - Implemented in FuzzyTech
- A better way would be a **theoretical justification** for a fuzzy operation along with their empirical confirmation.



“Debugging” and tuning MFs and operations

- A way to eliminate “blind” debugging is **obtaining ECCS instead of heuristically changing an operation.**
- In this approach, the debugging rules is needed only for correcting a linguistic rule table.
- Another way is **using interpolation based on picks of membership functions and values of the output variable for these picks.** I
- **Justification.**
- The common in fuzzy logic CoG defuzzification method produces an intuitively inconsistent wave in some output functions [Kovalerchuk, 1996b].
- The wave is inconsistent with the intuitive idea of sustained monotone growth of a trading environment indicator with improvement of interest rates and fees .
- The second interpolation that we present in here does not have a wave and it is much simpler than CoG for computation.

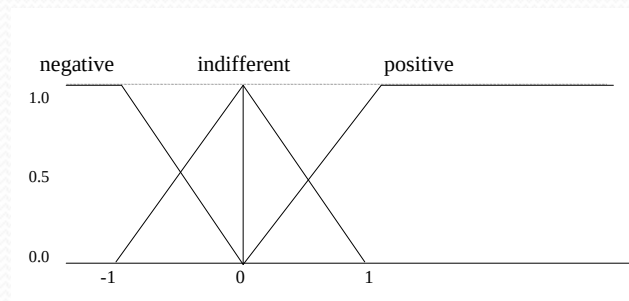
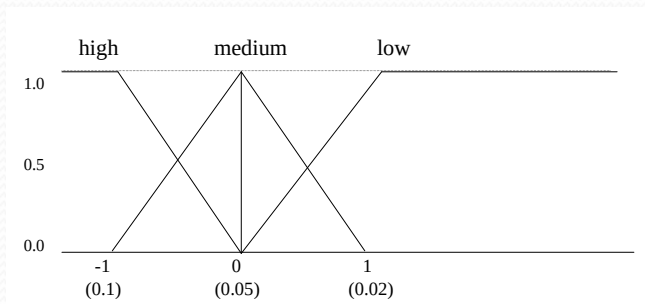


“Debugging” and tuning MFs and operations

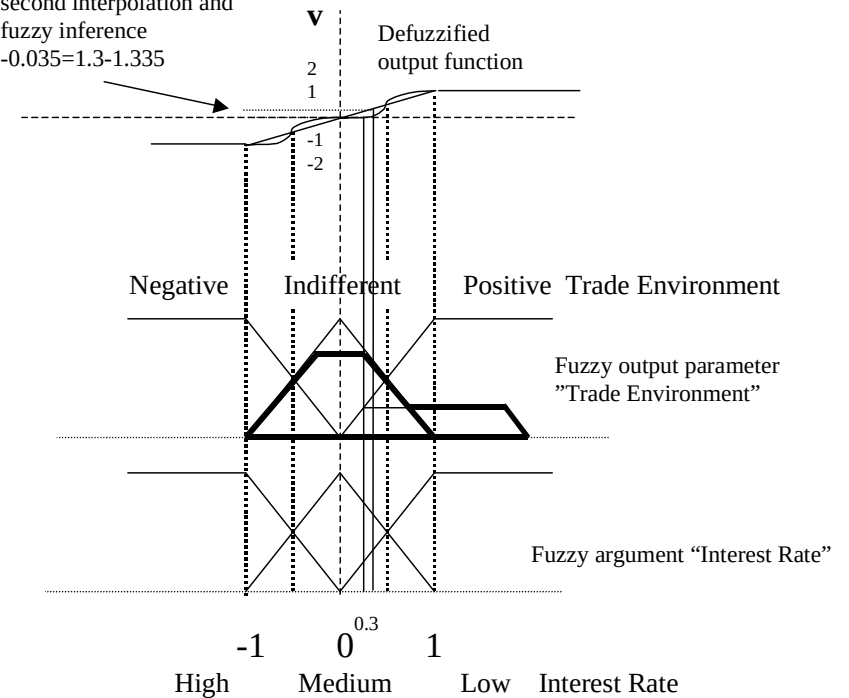
- **Algorithm.** The algorithm for debugging and tuning rules based on the second interpolation and ECCS consists of five procedures:
- Procedure 1. Identify Exact Complete Context Space (ECCS) with triangular membership functions
- Procedure 2. Construct symmetrical ECCS.
- Procedure 3. Identify monotone rules.
- Procedure 4. Transform ECCS to meet requirement of monotonicity.
- Procedure 5. Generate training data for pick points, 0.5-points and o-points of input MFs and linguistic variables in symmetric ECCS.
- Procedure 6. Test linguistic rules against training data.
- Procedure 7. Tune linguistic rules for matching training data.
- Procedure 8. Interpolate pick points of debugged output MFs of ECCS (see Figure 7.10).

“Debugging” and tuning MFs and operations

If-part		Then-part
Interest Rate	Trade Fee	Environment
1	1	1
-1	0	-1
0	0	0
1	0	1
0	1	(0)
-1	1	0



Difference between second interpolation and fuzzy inference
 $-0.035 = 1.3 - 1.335$



Advantages of "exact complete" context for fuzzy inference

- Advantages:
 - time and memory for input data,
 - simple Input-Output function,
 - complete set of "If-Then" Rules,
 - reliability of the output.
- Time and Memory for Input Data
- *Example.* Consider an input with linguistic terms: NL is "negative large"; NM is "negative medium"; NS is "negative small"; ZR is "approximately zero"; PS, PM, and PL are similarly abbreviated for the positive statements.
- Exact completeness of context.
$$NM(x) + NS(x) = 1.00.$$
- We can interpret $NL(x)$ and $NS(x)$ as "subjective probabilities" which form a probability space for a given x . Thus, the first advantage of this is that the same level of objectivity, stability, and context dependence is achieved here as in probability theory.
- The set of membership functions (MFs) represent all classes of such simple probability spaces.
- We say that such fuzzy sets form context (context space).
- Many linguistic variables have no this property, they represent incomplete context space ($\text{sum} < 1$) and over-complete spaces ($\text{sum} > 1$).



Context Space and MFs: compact representation for probabilities

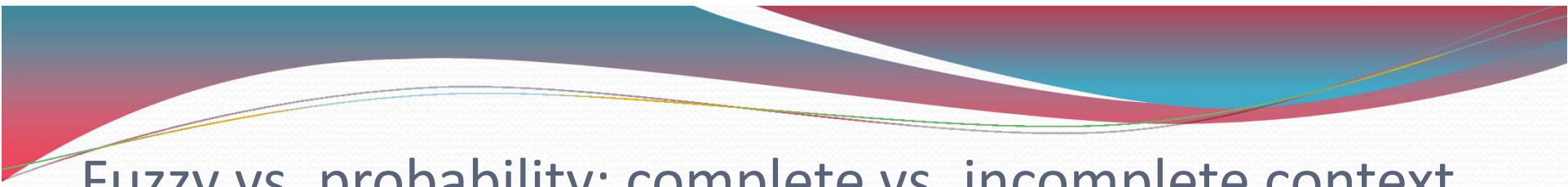
- **Compactness.** Let X be a set of 100 grades (elements); now, instead of 100 separate probability spaces with probabilities of two elementary events each, we simply use seven MFs.
 - In the first case, we would have to store at least 100 numbers. In the second case, we need store only seven numbers $(-3, -2, -1, 0, 1, 2, 3)$.
 - All other numbers are computed via linear functions based on these seven. Because of such compact representation, much less computer memory is needed.
 - We also can easily compute values of MFs for intermediate points, which are not included in our 100 points.
- **Time advantage.** The time saved by the expert who has to provide m values.
 - An expert would have to give answers about elementary probabilities for 200 events in the first case, but for the triangular MFs typically the expert needs to produce only seven numbers.
- While linguistic variables that are used in fuzzy inference can be represented by "over complete" or "incomplete" parameters, the majority of real fuzzy control systems are based on "**exact complete**" context space (ECCS).



Fuzzy vs. probability:

complete vs. incomplete context

- The knowledge of the agent could be captured by a single probability measure P on the frame of discernment [Dubois/Prade, 1998, 2001]
- BK:
- Only very limited knowledge of the agent could be captured by a single probability measure P on the frame of discernment.
- Human reasoning is contextual.
- Capturing the knowledge of the agent in context requires a set of probability measures P_i on a set of frames of discernment.

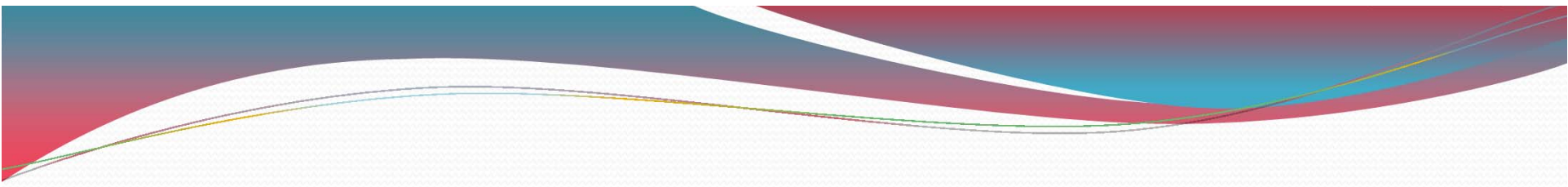


Fuzzy vs. probability: complete vs. incomplete context

- D&P:
 - Probabilistic knowledge is an extension of the complete knowledge case to many valued belief degrees, where the characteristic property of complete belief bases (K implies p or implies $\neg p$, for all p) is generalized by the property $P([p]) + P([\neg p]) = 1$.
 - A genuine extension of the Boolean incomplete knowledge situation to many-valued degrees of belief is obtained when the set E representing the agent knowledge is a fuzzy set, whose membership function then represents a possibility distribution π over the frame of discernment.
- BK:
 - Only very limited knowledge of the agent that is out of context could be captured by a single fuzzy set. Capturing the knowledge of the agent in context requires a set of fuzzy sets on a set of frames of discernment.
 - Probabilistic and fuzzy agent's knowledge representations in context are equivalent when captured by exact complete context spaces (ECCS).
 - The fuzzy sets representation of ECCS is more compact. Probabilistic agent's knowledge representation requires exact complete context spaces (ECCS). ECCS are common in fuzzy control.
 - Fuzzy agent's knowledge representation permits incomplete and over complete context spaces (ICCS, OCCS).
 - Agents knowledge representation in Exact Complete Context Space allows more rigorous reasoning.
 - ICCS and OCCS can be “debugged” to be transformed to ECCS.

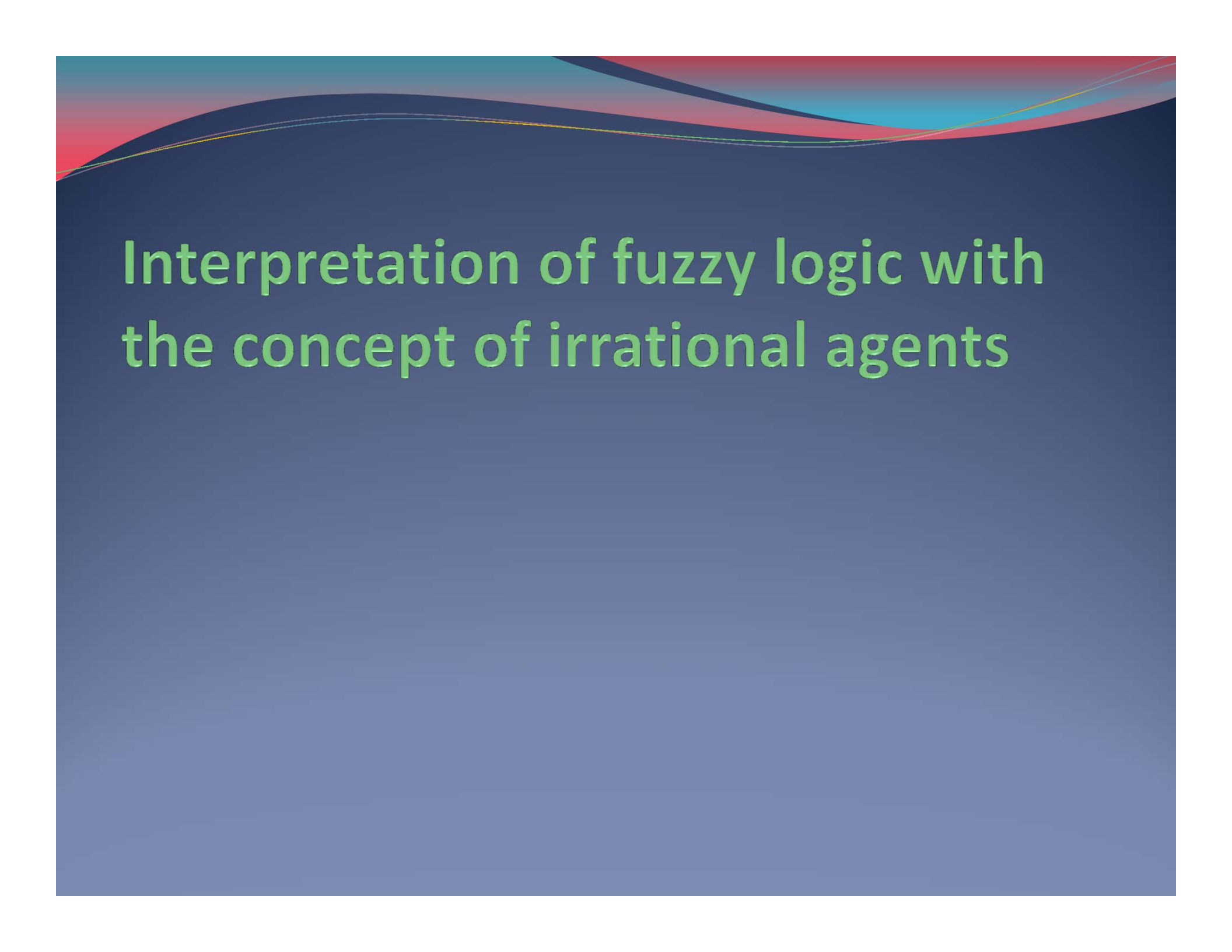
Fuzzy vs. probability: complete vs. incomplete context

- Example: Agent G produced two statements S_T and S_S :
 - S_T = “Probability that John, 175 is Tall is 0.7”, $P(G, 175, \text{Tall}) = 0.7$.
 - S_S = “Probability that John, 175 is Short is 0.4”, $P(G, 175, \text{Short}) = 0.4$.
- If “Short” is “not Tall” then we have a contradiction.
 $P([p]) + P([\neg p]) = 1$ implies that $P(G, 175, \text{Not Tall}) = 0.3$
- How to reconcile these statements? -- It can be done by clarifying context.
- Did agent G answered in both sentences in the same or different contexts of only two alternatives (Not Tall, Tall) or more alternatives, e.g., {Short, Medium, Tall}, {Very Short, Medium, Tall, Very Tall}.
- If $P(G, 175, \text{Tall}) = 0.7$ is within context {Short, Medium, Tall}, but $P(G, 175, \text{Short}) = 0.4$ within context {Short, Tall} then summation is not applicable.
- Consider agent's G statement $S_D = S_T \text{ OR } S_S$, Fuzzy logic will produce $m(S_D) = \max(0.7, 0.4) = 0.7$.
- Intuitively 0.7 is underestimate of this probability. How can we ‘debug’ it?
- We can use relations between two contexts. In the context {Short, Tall}, the “Short” category took some people from the “medium” category, inflating m to 0.4. By “debugging” it for the the context {Short, Medium, Tall} we can decrease it to $0.3 = m(G, 175, \text{Short})$. Then we can add 0.7 and 0.3 and get $m(S_D) = 1$ that is intuitively expected answer.
- The standard Fuzzy logic max operation for OR misses this ‘debug’ opportunity. Moreover it misses this opportunity even in the same context.

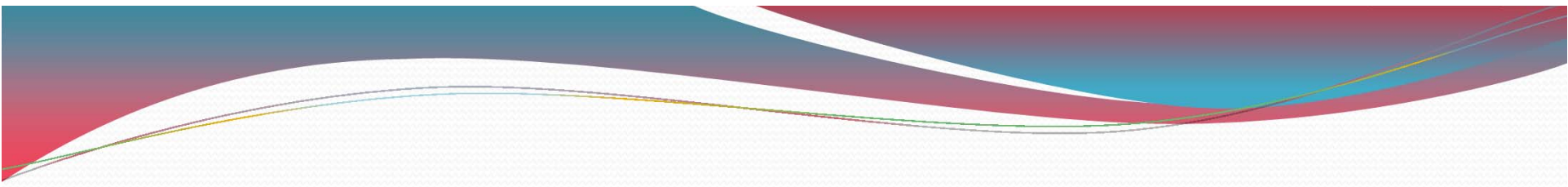


Possibility theory

- [Dubois, Prade, 1998, 2001]
- Let E be a set of possible states of fact. It is assumed that these states can be ordered in terms of plausibility, **normality**, and the like.
- The $\pi(\omega)$ reflects to what extent the state of affairs ω is unsurprising, expected as normal state of affairs.
- The overlapping between E and the ordinary set of models of a proposition p is a matter of degree.
- This degree of overlapping is the **degree of possibility** defined by $(p) = \max \pi(\omega), \omega \in [p]$
- The rationale: the plausibility of p is evaluated in the most normal situation where p is true.
- (p) is also the degree of consistency between the state of knowledge of the agent and the proposition p , and describes to what extent p is possibly true.
- Thus the possibility theory intends to estimate the normalcy of the situation where p is true, rather than the degree of truth of p itself.
- The idea is that highest normalcy is not necessary is the highest frequency. We may consider good weather normal (not unusual) but it may not be the most frequent weather.
- Alternatively normalcy can be computed probabilistically. Consider a statement “Probability the stock A will be between \$100-\$120 tomorrow is 0.7.” 0.7 can be frequency that stock A is between \$100 and \$120.
- The advantage of frequencies is that we have a operational way to get an estimate of normalcy.

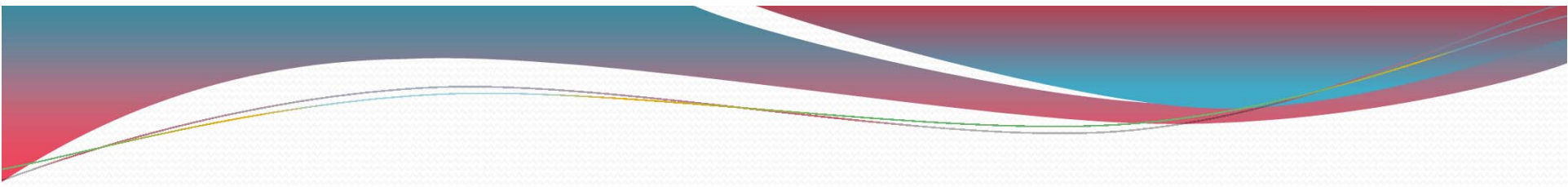


Interpretation of fuzzy logic with the concept of irrational agents



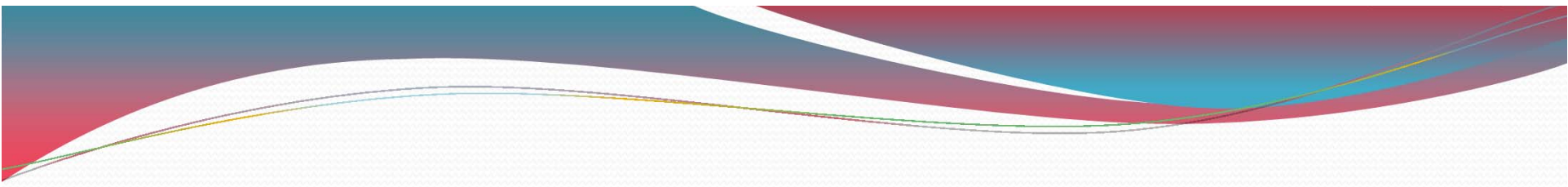
Assigning input truth values

- Modern axiomatic uncertainty theories (fuzzy logic, probability theory and others) provide a calculus for manipulating with probabilities, membership functions, degrees of belief and other uncertainties when the values such as uncertainties of input variables are already given.
 - These theories do not include a common mechanism for getting the input uncertainties. The value of these theories is in computing uncertainties of complex events that follow a structure imposed by axioms of a specific uncertainty theory.
 - Similarly Agent Logics such as Normative Logic, Pragmatic Logic and Epistemic Logic can model set of agents, but again leave assigning input truth values outside of these logics.
- The **lack of internal mechanisms** for getting input uncertainty values often means in the end that the same mechanism (based on frequencies or subjective probabilities) is applied for getting input probability values, fuzzy logic membership functions, and belief functions.
 - This is a source of much confusion -- what is the real difference between all of these theories?
 - Why are their axioms different if input uncertainty values are computed essentially in the same way? A resolution of this confusion is critical from both theoretical and practical viewpoints.
- We show that a common internal mechanism of getting uncertainty values can be built by incorporating *rational* and *irrational*, *conflicting* and *interacting agents* and worlds along with their internal and external contexts.



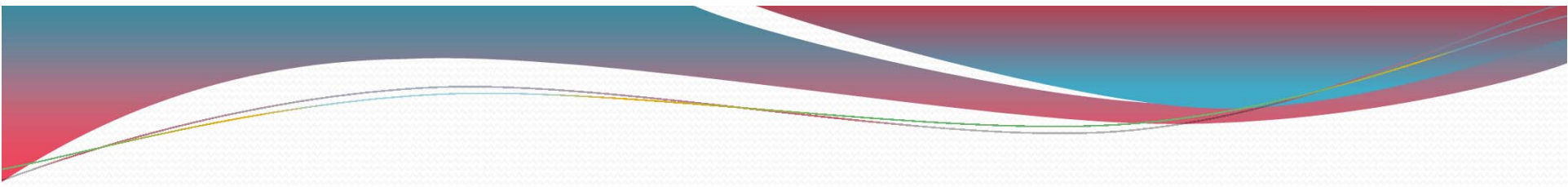
Agent Logic of Uncertainty (ALU)

- The probability theory computes the uncertainty of complex events that model the behavior of a society of **rational agents**. Real agents can be rational only at some extent and be quite *irrational* in many aspects.
- This leads to a fundamental difference in the way in which theories of uncertainties can justify their axioms. Some theories (logics) of uncertainties can be set up mathematically with or without experiments.
- We call logics that are justified without direct experiments **prescriptive theories (logics)** of uncertainty. They tell how agent should assign uncertainty values to be rational.
- Alternative agent logics of uncertainty are dependent on empirical, physical knowledge, and experiments about real behavior of agents in specific contexts. These logics are **descriptive** -- they describe how real agents actually combine uncertainty values. Such logics cannot be built by pure mathematical axiomatic means, they need experiments or experience from the open real world.
- We argue **that fuzzy logic should be built as such descriptive theory** that aggregates uncertainty values produced by a **mix of rational, irrational, conflicting or inconsistent agents**. We call this area the Agent Logic of Uncertainty (ALU) [Kovalerchuk, Resconi, 2008-2011] and describe in the tutorial.



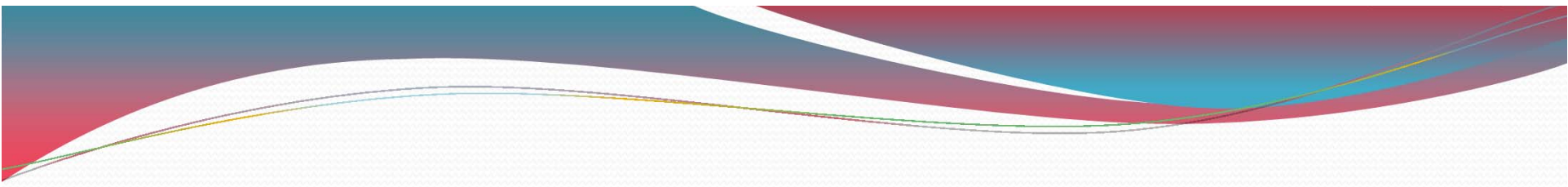
Conclusion

- For correct inference under linguistic uncertainty in applications, it is very useful and necessary to construct exact complete context spaces.
- Unless one specifies the appropriate, necessary context space in which s/he is working, correct inferential solutions cannot be clearly arrived at, nor can one establish a clear foundation upon which to debate the efficacy of the problems under consideration.
- Further development of the context spaces should deal with additional requirements on the components of specifiable context spaces.



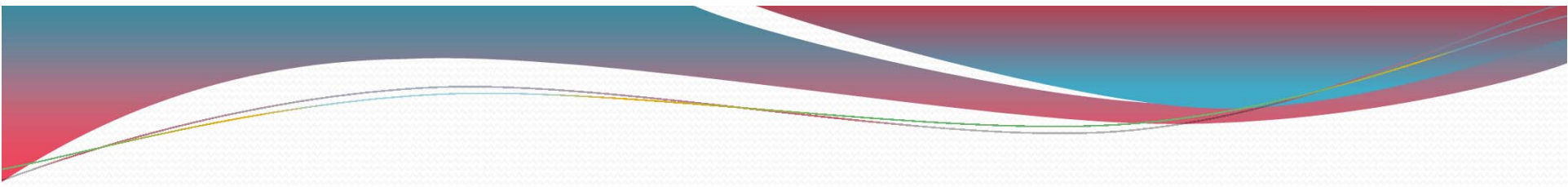
Conclusion

- Zadeh's fuzzy linguistic variables embedded into the exact complete context spaces have *fundamental representational advantage vs. multiple equivalent small probability spaces*.
- In other words, *few fuzzy membership functions* in a *linguistic variable* provide a *quick way* to build a huge set of simple probability spaces.
- In this sense, fuzzy membership functions and probabilities are complimentary not contradictory. Thus, they are mutually beneficial by combining fast model development and rigor.



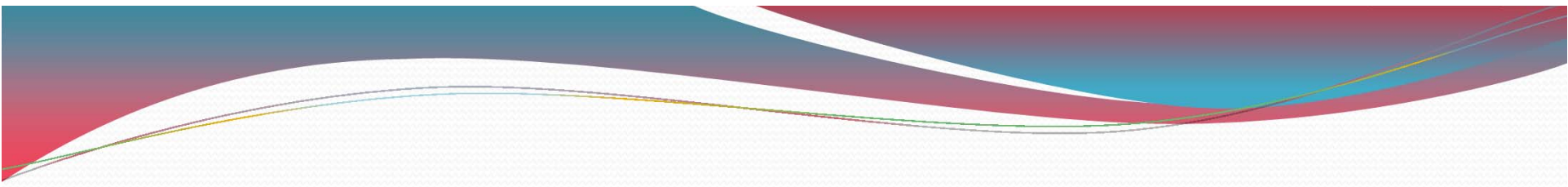
The tasks for the future

- Formulate more CWW tasks in terms of Context Spaces in both probabilistic and fuzzy logic terms
- Finding appropriate operations in the specific examples when both formulations are possible.



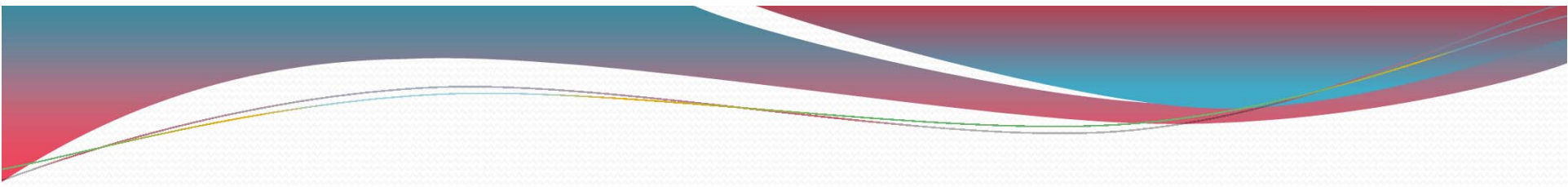
Bio

- Dr. Kovalerchuk is a professor of Computer Science at Central Washington University, USA. He published two books on Data Mining and Visual and Spatial Analysis (Springer 2000, 2005), 10 book chapters, and over 150 research papers. Dr. Kovalerchuk received his Ph.D dual degree in Computer Science and Applied Mathematics in 1977 from the Russian Academy of Sciences and his MS degree in Math at the Novosibirsk University (Russia). During the past 30+ years, he worked for both Industry and Academia. Dr. Kovalerchuk has worked and led multiple projects in the areas of Artificial Intelligence, Pattern Recognition, Image and Signal Processing, Data Mining and Fusion. His research has been funded by several US Federal Agencies. Recently Dr. Kovalerchuk chaired two International Conferences in Computational Intelligence. He also served as a Senior Visiting Scholar at the US Air Force Research Laboratory (AFRL). As a recognized expert in the field of Computational Intelligence, he has been invited and served as expert at US Government Panels and delivered multiple tutorials on these subjects.



Selected Publications

- Kovalerchuk B., Vityaev, E., Data Mining in Finance: Advances in Relational and Hybrid Methods, Kluwer Academic Publishers 2000, 325 p.
- Kovalerchuk B., Perlovsky, L., Wheeler, G., "Modeling of Phenomena and Dynamic Logic of Phenomena," Journal of Applied Non-classical Logics, 22(1): 51-82. <http://philsci-archive.pitt.edu/8785/1/DLP-draft.pdf>
- Kovalerchuk B., Perlovsky L. Uncertainty Modeling for Spatial Data Fusion and Mining, In: Proc. of 2011 IEEE Symposium Series on Computational Intelligence (SSCI), Paris, France, 11 Apr-15 Apr 2011
- Kovalerchuk B., Resconi., G. Agent-based Uncertainty Logic Network, 2010 IEEE World Congress on Computational Intelligence, Barcelona, Jul 18-23, 2010.
- Kovalerchuk B., Interpretable Fuzzy Systems: Analysis of T-norm interpretability, 2010 IEEE World Congress on Computational Intelligence, Barcelona, Jul 18-23, 2010.
- Resconi, R., Kovalerchuk, B., Agents' model of uncertainty, Knowledge and Information Systems Journal, Springer London, vol. 18, no. 2, pp. 213-229, Feb 2009, DOI 10.1007/s10115-008-0164-0
- Kovalerchuk, B. [Context spaces as necessary frames for correct approximate reasoning](#). *International Journal of General Systems*, v.25, n 1, 1996, pp. 61-80.
- Kovalerchuk, B., Klir, G. Linguistic context spaces and modal logic for approximate reasoning and fuzzy-probability comparison. In: *Proc. of Third International Symposium on Uncertainty Modeling and Analysis and NAFIPS' 95*, College Park, Maryland, 1995, IEEE Computer Society Press, 1995, pp.A23-A28.
- Kovalerchuk, B., Talianski, V. Comparison of empirical and computed values of fuzzy conjunction. *Fuzzy sets and systems*, v. 46, 1992, pp. 49-53



Selected Publications

- Kovalerchuk, B. Spaces of Linguistic Contexts: Concepts and Examples. In: *Second European Congress on Intelligent Techniques and Soft Computing*, Aachen, Germany, 1994, pp. 345-349.
- Kovalerchuk, B., Shapiro, D. On the relation of the probability theory and the fuzzy sets foundations. *Computers and Artificial Intelligence*, v.7, 1988, pp. 385-396.
- Kovalerchuk B. Review of fuzzy optimization foundations. 4th World Congress of the International Fuzzy Systems Association (IFSA'91), Brussels, 7-12 July 1991, Computer Management and Systems Sci. D. Dubois, R. Lowen, M. Roubens (Eds), Brussels, 1991, pp.129-132.
- Kovalerchuk B. Analysis of Gaines's logic of uncertainty. A Quarter Century of Fuzziness, Approximate Reasoning and Knowledge Engineering, Proc. of NAFIPS' 90 conference, Toronto, June 6-8, 1990. I.V.Turksen (Ed.), v. 2. pp. 293-295.
- B. Kovalerchuk, Probability vs. Fuzziness/ Zadeh <http://mybisc.blogspot.com/2010/12/probability-vs-fuzziness-zadeh.html>
- D. Dubois, H. Prade, Possibility theory: qualitative and quantitative aspects, In: Handbook of defeasible reasoning and uncertainty management systems, v. 1, Eds. D. Gabbay, and P. Smets, pp. 169-226., 1998
- D. Dubois, H. Prade, Possibility theory, probability theory and multiple-valued logics: A clarification, *Annals of Mathematics and Artificial Intelligence* 32: 35-66, 2001.
http://julien.chauveau.free.fr/m2reccid/MA1/assets/Dubois_PossibilityTheory.pdf