

Uncertainty for Qualitative Variables

SAND2007- XXXX

Presented to Consequence Management Team

November 12, 2007

John Darby

6462

Sandia National Laboratories



Acknowledgments

- **Approximate Reasoning based on Logic Evolved Decision Techniques developed at LANL by Terry Bott and Steve Eisenhower**
- **Assistance with Mathematics from Jon Helton, Arizona State/SNL**
- **Belief/Plausibility for Fuzzy Sets for Numeric Variables: Ron Yager paper**
- **Adversary Defender Model Concepts**
 - **Peter Merkle, SNL**
 - **Brian Bush, LANL**
- **Graphical Ranking Technique**
 - **John Cummings, SNL**



References

- “Evaluation of Risk from Acts of Terrorism: The Adversary/Defender Model using Belief and Fuzzy Sets”, SAND2006-5777, September, 2006
- “Evaluation of Risk for Acts of Terrorism using Belief and Fuzzy Sets”, Journal of Nuclear Materials Management, Winter, 2007, Volume XXXV, Number 2
- “LinguisticBelief: A Java Application for Linguistic Evaluation using Belief, Fuzzy Sets, and Approximate Reasoning”, SAND2007-1299, March, 2007
- “Linguistic Evaluation of Terrorist Scenarios: Example Application”, SAND2007-1301, March, 2007



Presentation Topics

- **Methodology**
- **Software Tools**



Methodology

- **Uncertainty**
 - **Ambiguity**
 - **Belief/Plausibility Measure**
 - **Superset of Probability Measure**
 - **Vagueness**
 - **Fuzzy Sets**
 - **Typically used for Numeric Variables**
- **Qualitative Variables**
 - **Fuzzy Sets**
 - **Purely Linguistic Fuzzy Sets**
- **Combination of Qualitative Variables**
 - **Convolution of Evidence over Linguistic Fuzzy Sets**
 - **Approximate Reasoning**
 - **Belief/Plausibility Measure for Uncertainty**



Uncertainty

- **Ambiguity**

- Uncertainty as to what will occur **in the future**
 - Dow Jones Industrial Average Close on Dec. 31, 2007
 - Will be one value
 - Ambiguity as to what that value will be

- **Vagueness**

- Uncertainty as how to categorize a **known** outcome
 - Dow Jones close is 13,876 on Dec. 31, 2007
 - Is this “High” ?
 - What do you mean by “High”?
- Vagueness can be expressed Linguistically (Words)



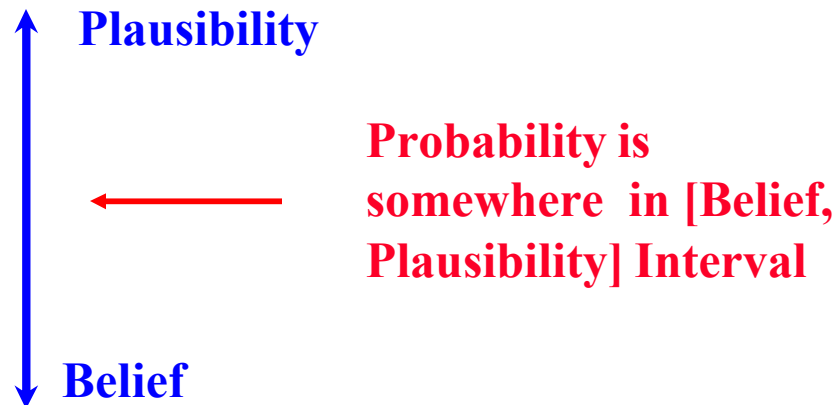
Ambiguity: Aleatory and Epistemic

- For a Fair Coin
 - Uncertainty is **Aleatory** (random)
 - Probability Heads is $\frac{1}{2}$
 - Probability Tails is $\frac{1}{2}$
- But if we cannot toss coin, we do not know coin is fair, we do not even know if coin has Heads and Tails
 - May not be Fair Coin (may be Weighted for Tails)
 - May be Two-Headed or Two-Tailed Coin
 - **Epistemic** (state of knowledge) uncertainty
 - Insufficient information to assign Probability to Heads and Tails
 - For Total Ignorance
 - Belief/Plausibility for Heads is 0/1
 - Belief/Plausibility for Tails is 0/1
- With more information (actually tossing the coin) we can reduce Epistemic Uncertainty
 - If at least one Heads and one Tails occur in a series of tosses, we know coin has Heads and Tails
 - Many tosses needed to assess if coin is fair
- For Fair Coin we cannot reduce aleatory uncertainty



Belief and Plausibility

- Belief / Plausibility form a Lower / Upper Bound for Probability
- Belief is what probability **will** be
- Plausibility is what probability **could** be

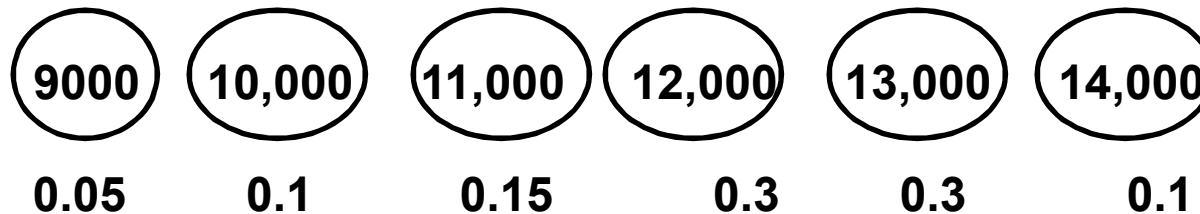


- Similar to a Confidence Interval for a Parameter of a probability distribution; a confidence measure that parameter is in interval, but exactly where in interval is not known
- Belief/Plausibility both reduce to Probability if Evidence is Specific

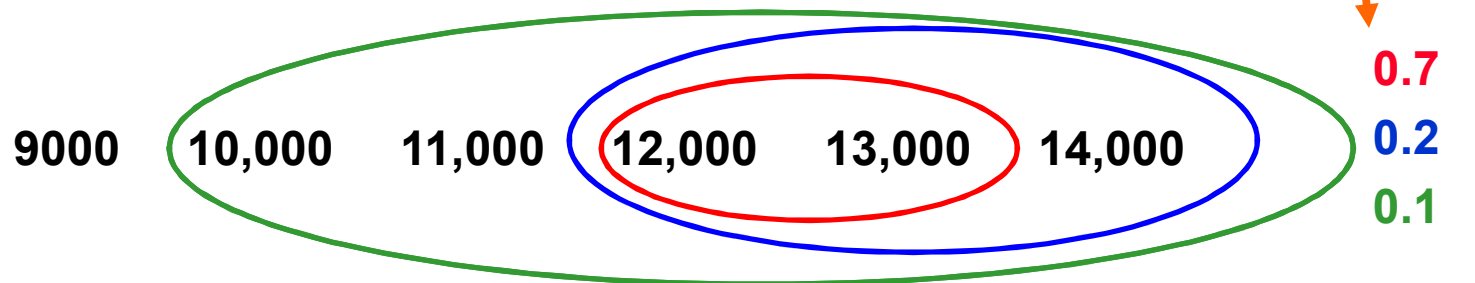


Example of Evidence: Predict Stock Market Close Dec. 31, 2007

- Probability



- Belief/Plausibility



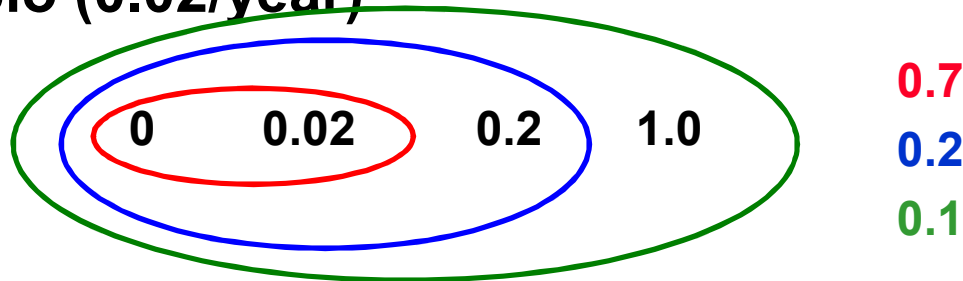


Why is this useful?

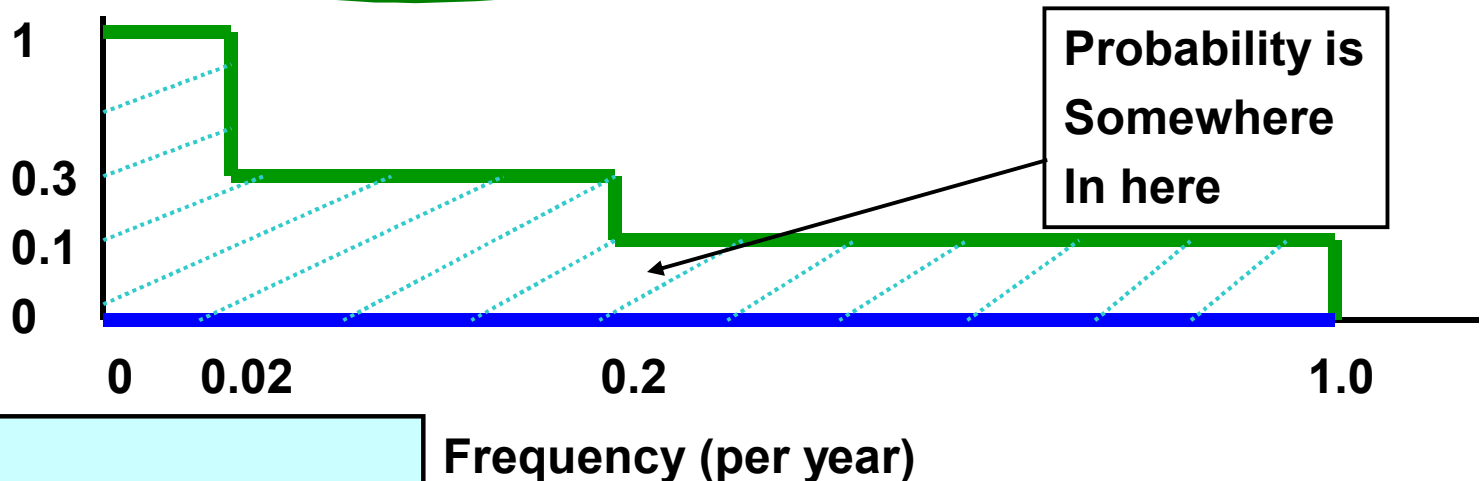
- What is Likelihood of Bio-Terror attack against a Major US City?
 - Frequency of Attack (per year) somewhere in $[0, 1]$
 - Point Estimate? **Useless**, too much uncertainty
 - Probability Distribution?
 - Must assign probability to each value in $[0, 1]$
 - Probability for 0 attacks per year is high
 - Probability assigned to 0 affects probabilities for all other values as **must sum to 1.0**
 - Is $P(0) = 0.01, 0.001, 0.00001, \dots$? **Don't Know**

Why is this useful?

- Evidence is about 1 major Attack every 5 years (0.2/year)
 - Assume Expert Opinion is: 10% Chance Attack is bio (0.02/year)



Likelihood
Exceed
Frequency:
Plausibility
Belief



Automated in
BeliefConvolution Java code

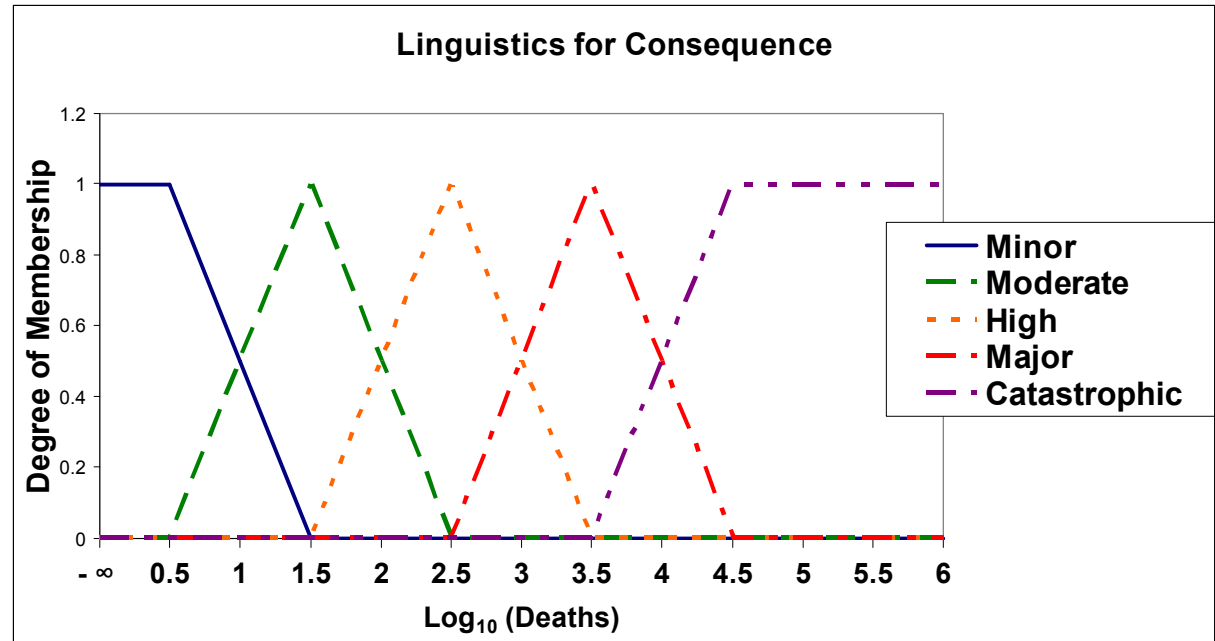
Frequency (per year)

jldarby@sandia.gov

Vagueness

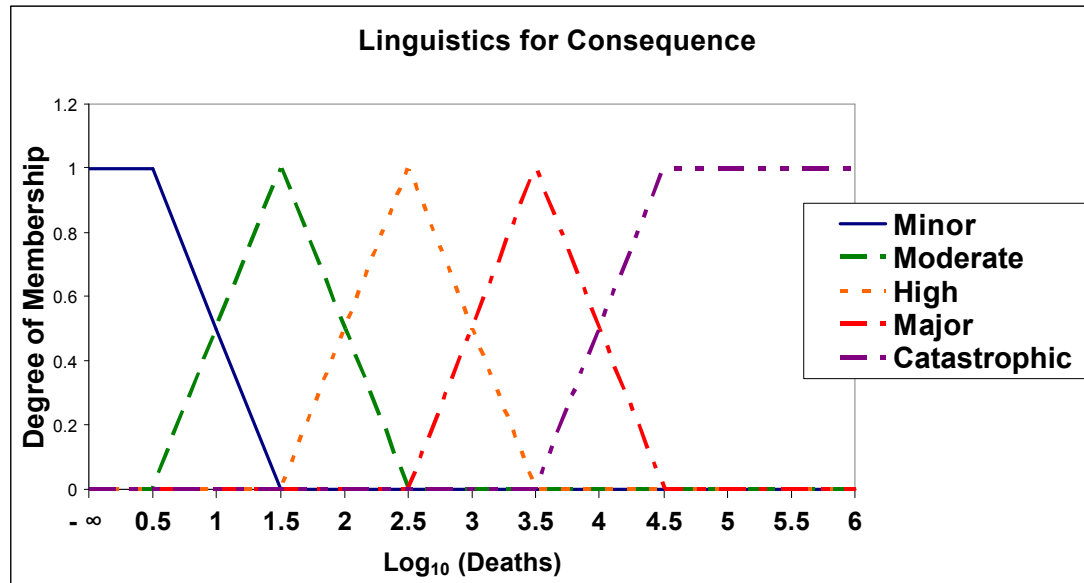
Fuzzy Sets for Numeric Variable

- Represent Variable with **Sets** to **reason at Fidelity Desired**. Above 30,000 deaths is “Catastrophic”.
- Use **Fuzzy Sets** to **Avoid Sharp Distinction**. “Major” Deaths is Between *About* 1000 and *About* 10,000. 999 and 1001 deaths are each part “High” and part “Major”.

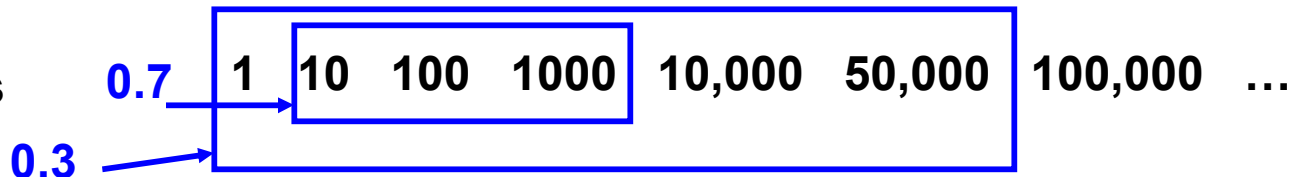


Uncertainty for Fuzzy Sets: Numeric Variable

Fuzzy Sets for Deaths



Evidence For Deaths



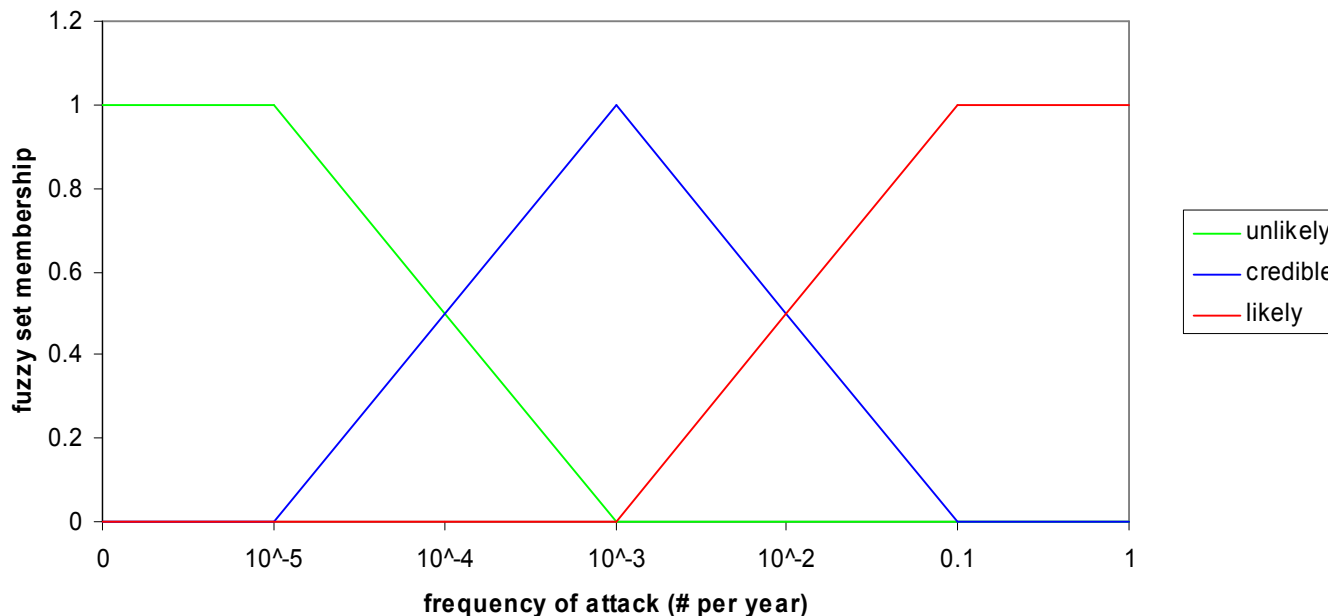
Uncertainty Distribution

| for Deaths: | Minor | Moderate | High | Major | Catastrophic |
|-----------------------|----------|----------|-------|----------|--------------|
| Belief / Plausibility | 0 / 0.65 | 0 / 1 | 0 / 1 | 0 / 0.65 | 0 / 0.3 |

Why is this Important?

- What is Likelihood of Bio-Terror attack against a Major US City
 - Evidence is about 1 major Attack every 5 years (0.2/year)
 - Assume Expert Opinion is: 10% Chance Attack is bio (0.02/year)
- Assume Following Fuzzy Sets for Evaluating Frequency of Attack

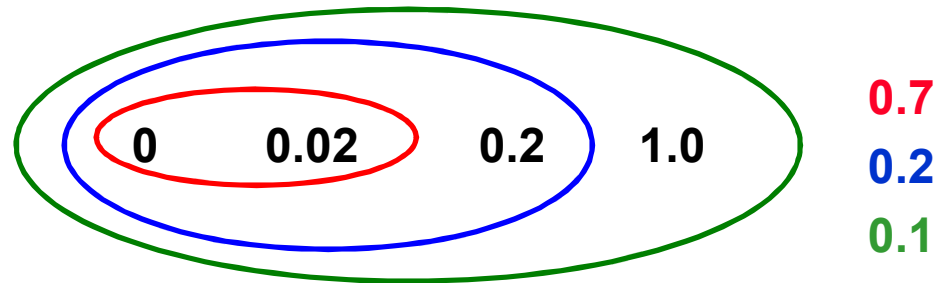
Defender Fuzzy Sets for Threat



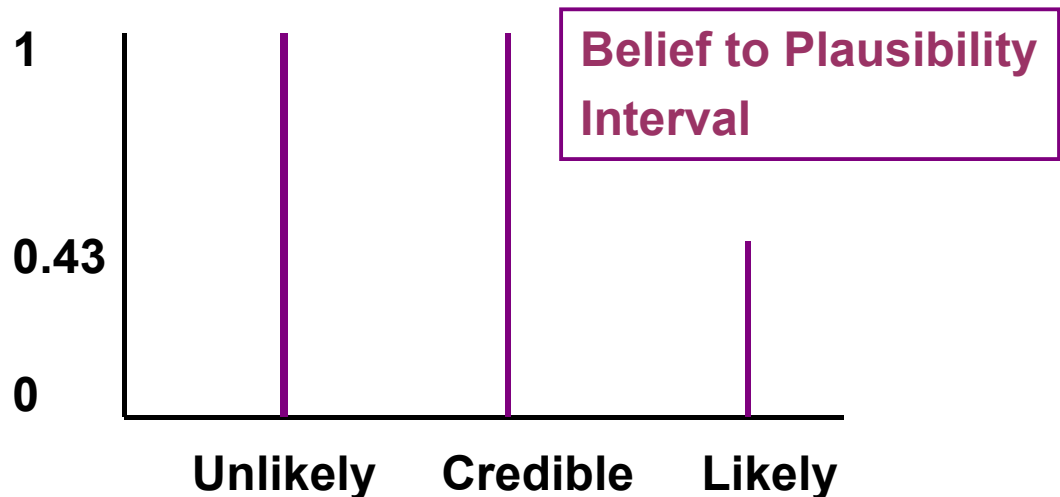


Why is this Important?

Evidence



Likelihood
For Frequency



*(Calculated with BeliefConvolution
code using Yager Method)*



Qualitative Variables

- Variable Segregated into **Purely Linguistic Fuzzy Sets**
 - Variable: “Health”
 - Fuzzy Sets: “Bad”, “Moderate”, “Excellent”
- Why Pure Linguistics?
 - **Numeric Scale is Unknown**
 - Is “Health” $[0, 10]$, $[0, 10^6]$, $[-700, \text{square root of } 42]$?
 - **Scaling is Un-Manageable when Combine Variables**
 - Combine “Health” with “Wealth” to Evaluate “Quality of Life”
 - “Wealth” can be Numeric: $[\$0, \$50\text{B}]$
 - What is Numeric Scale for “Health”?
 - What is Numeric Scale for “Quality of Life”?



Fuzzy Sets for Non-Numeric Variable

Adversary Level of Technical Training:

High School

Bachelors

Advanced

Do **NOT** Force Numeric Measure: Requires Arbitrary Scale

Adversary Level of Technical Training:

High School = 1?

Bachelors = 2?

Advanced = 3?

Adversary Level of Technical Training:

High School = 10?

Bachelors = 100?

Advanced = 1000?

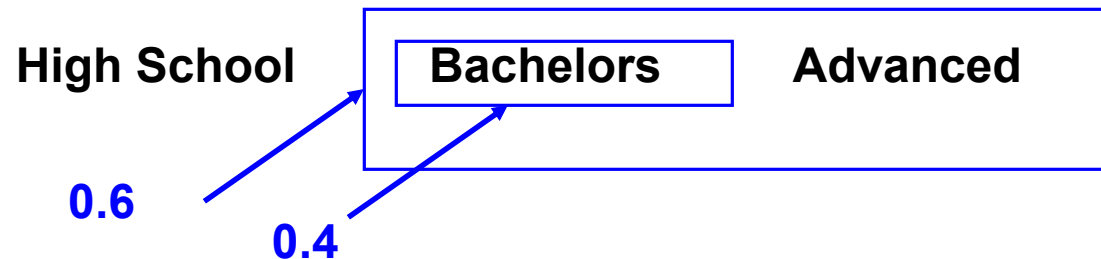


Uncertainty for Fuzzy Sets: Non-Numeric Variable

- **Fuzzy Sets for Adversary Level of Technical Training**

– High School Bachelors Advanced

- **Evidence**



- **Uncertainty Distribution: Belief / Plausibility**

| High School | Bachelors | Advanced |
|-------------|-----------|----------|
| 0 / 0 | 0.4 / 1 | 0 / 0.6 |



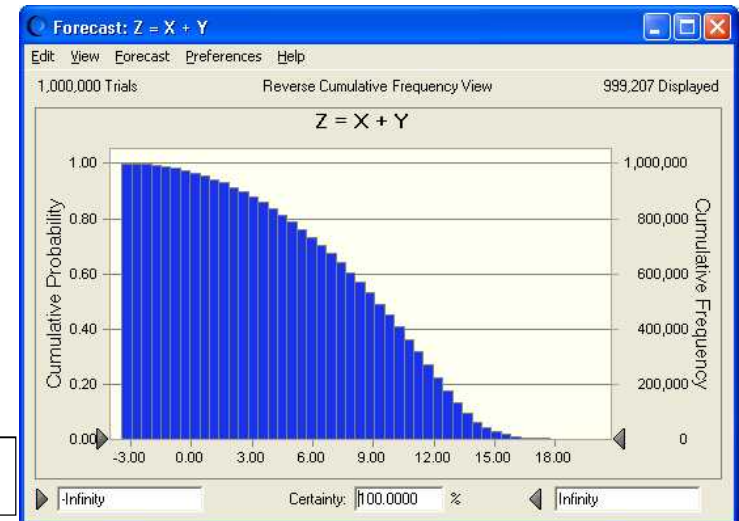
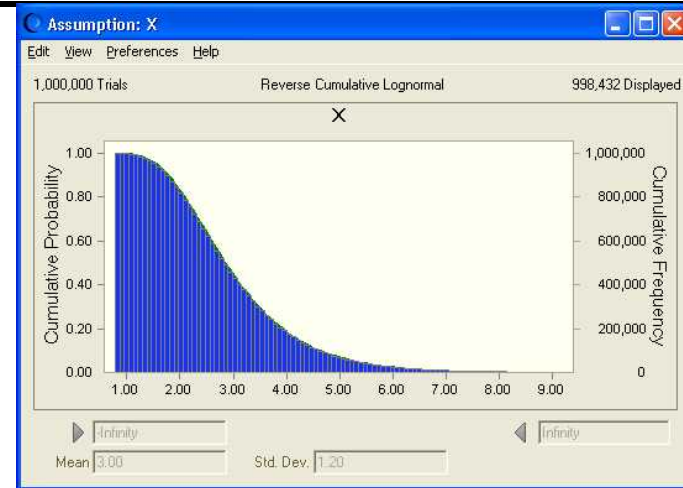
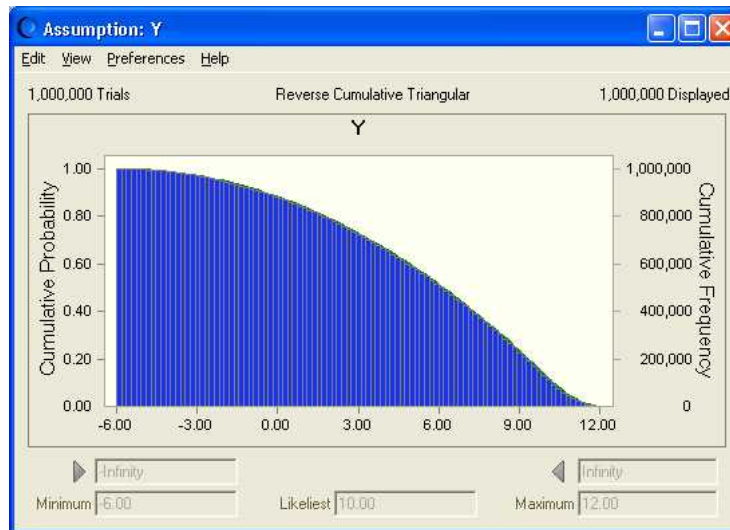
Combining Variables: Convolution of Uncertainty Distributions

- **Belief/Plausibility Distributions**
 - Evidence Over Fuzzy Sets for Each Variable
- **Convolute Distributions** per the Rule Base
 - Mathematics of Belief/Plausibility
- **Same Concept as Convolution of Probability Distributions**
 - Mathematics of Probability

Convolute Probability Distributions: Crystal Ball Software

$$Z = X + Y$$

X and Y Independent



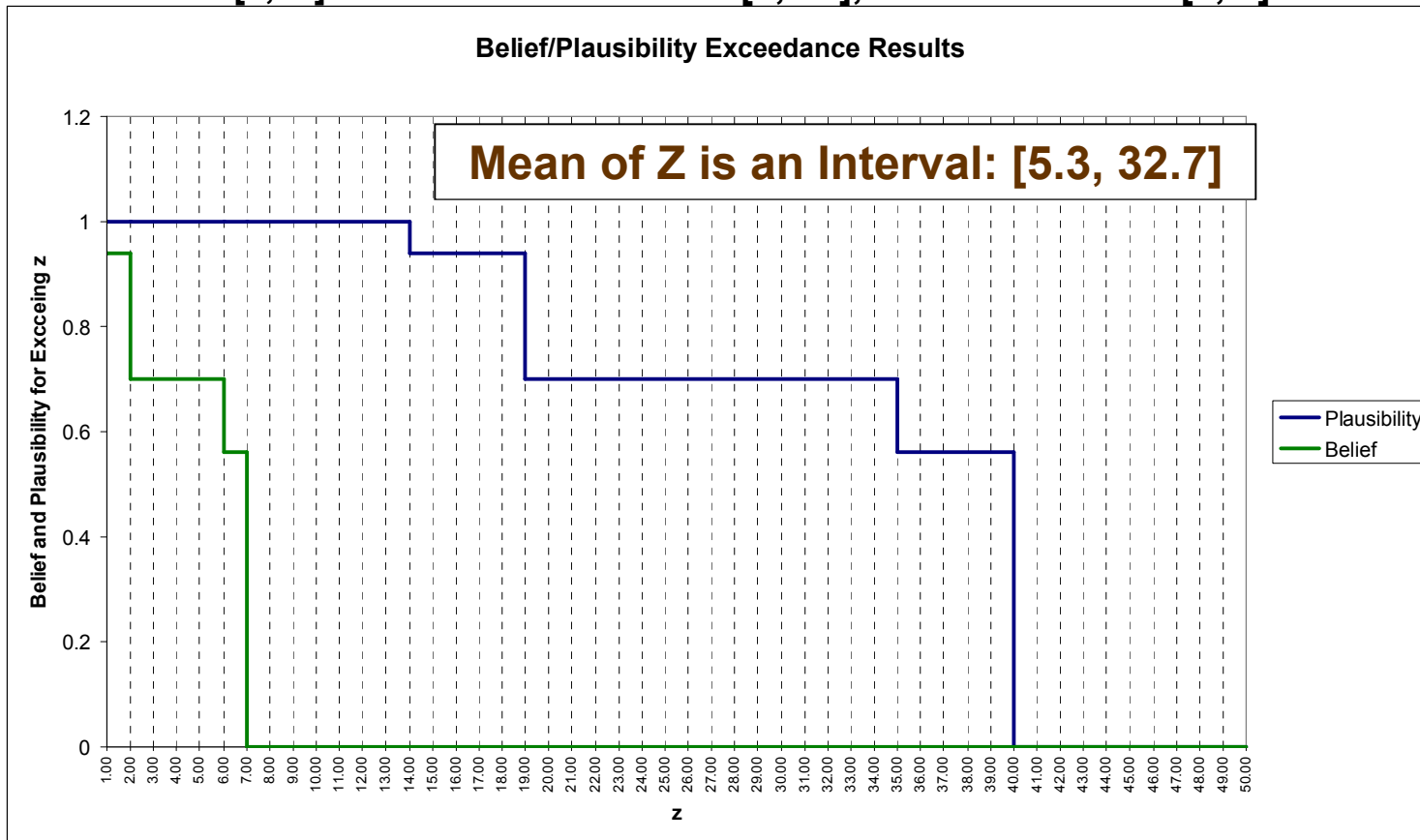
Mean of Z is a Point value: 8.34

Convolute Belief/Plausibility Distributions for Numeric Variables: BeliefConvolution Software

$Z = X + Y$, X and Y Non-Interactive

X over $[1, 20]$ with Evidence: 0.8 for $[2, 15]$, 0.2 for $[1, 10]$

Y over $[0, 30]$ with Evidence: 0.7 for $[5, 25]$, Evidence 0.3 for $[0, 4]$





Convolute Belief/Plausibility Distributions for Linguistic Variables: LinguisticBelief Software

- **Example Follows**



Combining Qualitative Variables: Approximate Reasoning

- Mathematics for **Combining Words**
- If we use Words instead of numbers we need a way of combining the Words for Different Variables
- Implemented as A **Rule Base for Combining Fuzzy Sets from Different Variables**



Combination of Linguistic Variables: Example

Develop the Model: Happiness for Any Individual
Define the Variables and their Fuzzy Sets

- **Basic Variables**
 - **Health**
 - Bad, Moderate, Excellent
 - **Wealth**
 - Poor, Middle Class, Rich
 - **Outlook on Life**
 - Pessimist, Optimist
- **Rule Based Variables**
 - **Quality of Life = Health x Wealth (x per rule base)**
 - Not so Good, Good
 - **Happiness = Outlook on Life x Quality of Life**
 - Depressed, Accepting, Very Happy

Combination of Linguistic Variables: Example

**Develop the Approximate Reasoning Rule Base for
Rule Based Variables**

Quality of Life

Rules for selected RuleLinguistic

Rules for RuleLinguistic: Quality of Life

| Fuzzy Set for Input Linguistic: Health | Fuzzy Set for Input Linguistic: Wealth | Output Fuzzy Set for Rule (blank if rule not set) |
|--|--|---|
| Bad | Poor | Not so Good |
| Bad | Middle Class | Not so Good |
| Bad | Rich | Not so Good |
| Moderate | Poor | Not so Good |
| Moderate | Middle Class | Not so Good |
| Moderate | Rich | Good |
| Excellent | Poor | Good |
| Excellent | Middle Class | Good |
| Excellent | Rich | Good |

Specify Output Fuzzy Set for Selected Rule Choices Are: ▼

Accept Rules as Shown Cancel

Combination of Linguistic Variables: Example

Happiness

Rules for selected RuleLinguistic

Rules for RuleLinguistic: Happiness

| Fuzzy Set for Input Linguistic: Outlook on Life | Fuzzy Set for Input Linguistic: Quality of Life | Output Fuzzy Set for Rule (blank if rule not set) |
|---|---|---|
| Pessimist | Good | Accepting |
| Pessimist | Not so Good | Depressed |
| Optimist | Good | Very Happy |
| Optimist | Not so Good | Accepting |

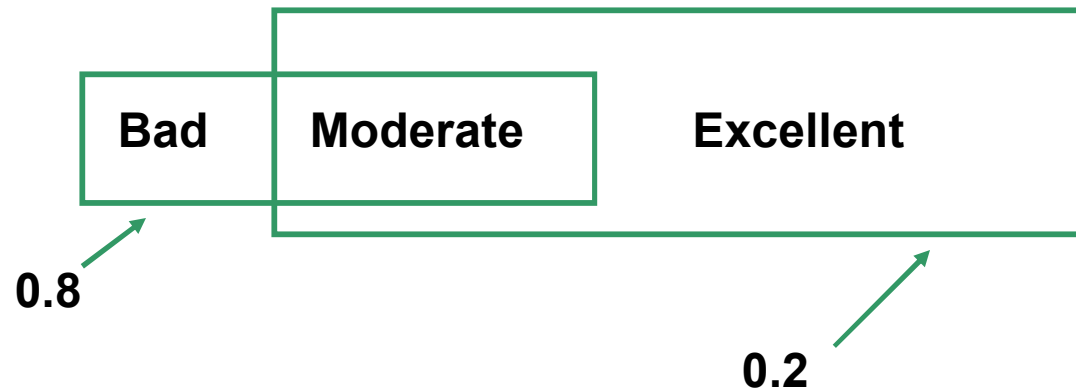
Specify Output Fuzzy Set for Selected Rule Choices Are: ▼

Accept Rules as Shown Cancel

Combination of Linguistic Variables: Example

**Evaluate the Model for Specific Individual:
Happiness for “John”**

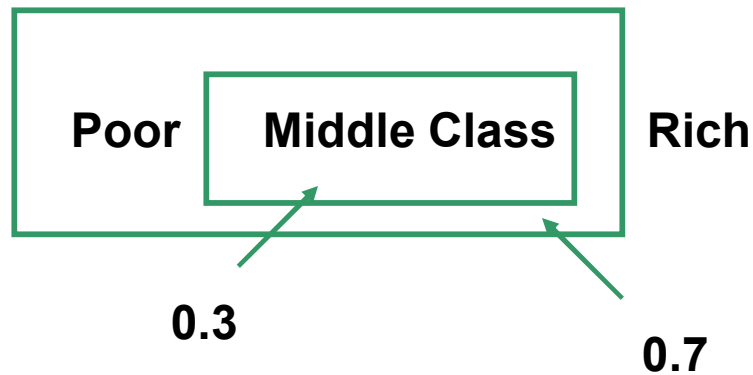
Assign Evidence to Fuzzy Sets for Basic Variables
Health





Combination of Linguistic Variables: Example

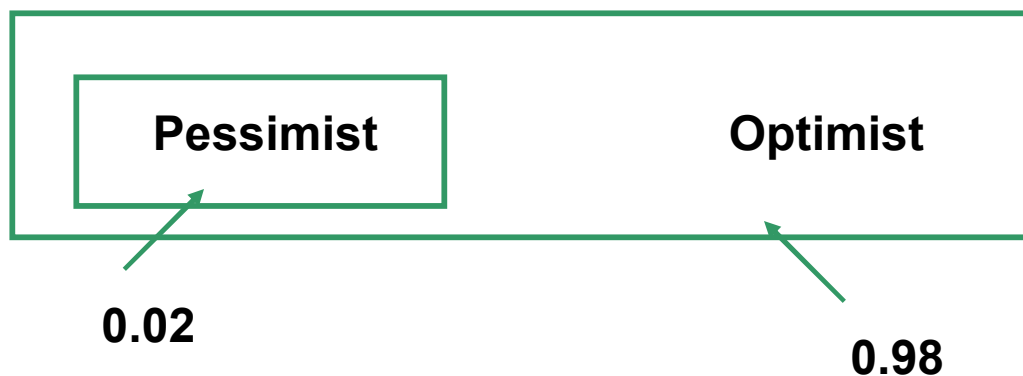
Wealth





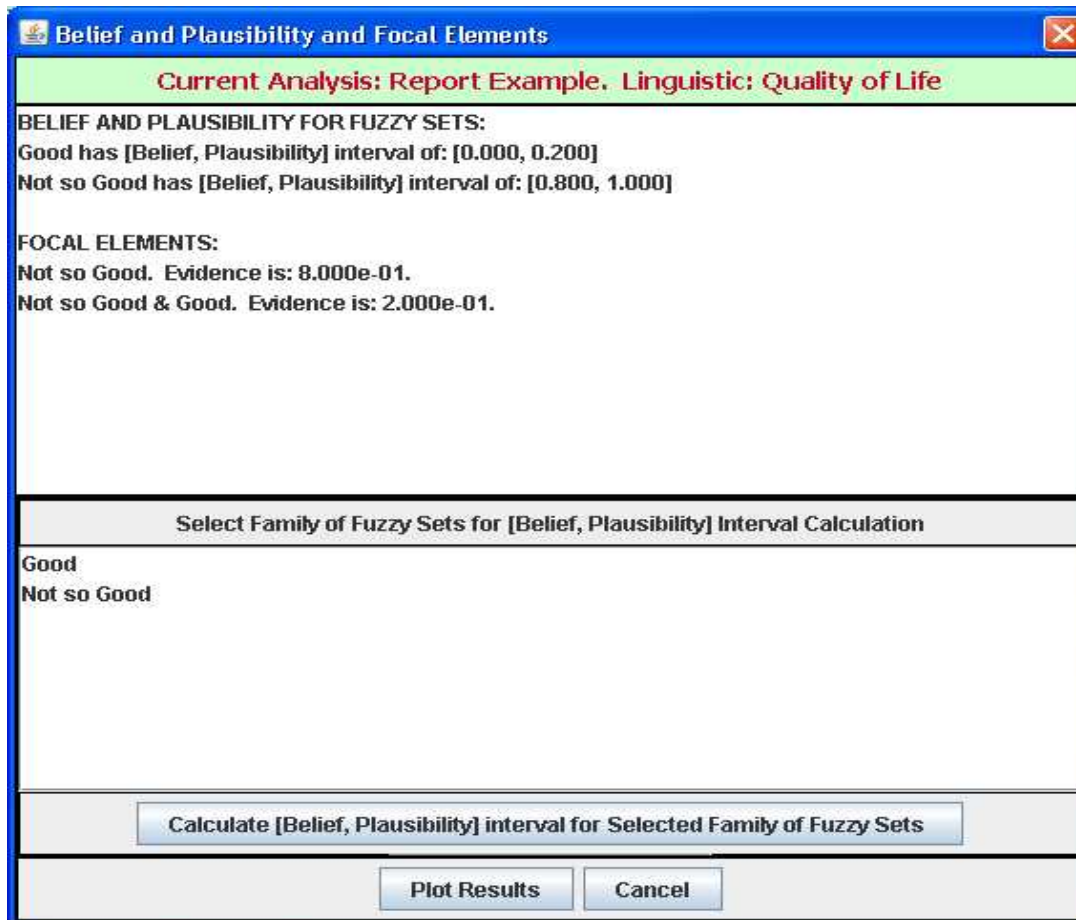
Combination of Linguistic Variables: Example

Outlook on Life



Combination of Linguistic Variables: Example

Evaluate Variable: Quality of Life for John



Belief and Plausibility and Focal Elements

Current Analysis: Report Example. Linguistic: Quality of Life

BELIEF AND PLAUSIBILITY FOR FUZZY SETS:
Good has [Belief, Plausibility] interval of: [0.000, 0.200]
Not so Good has [Belief, Plausibility] interval of: [0.800, 1.000]

FOCAL ELEMENTS:
Not so Good. Evidence is: 8.000e-01.
Not so Good & Good. Evidence is: 2.000e-01.

Select Family of Fuzzy Sets for [Belief, Plausibility] Interval Calculation

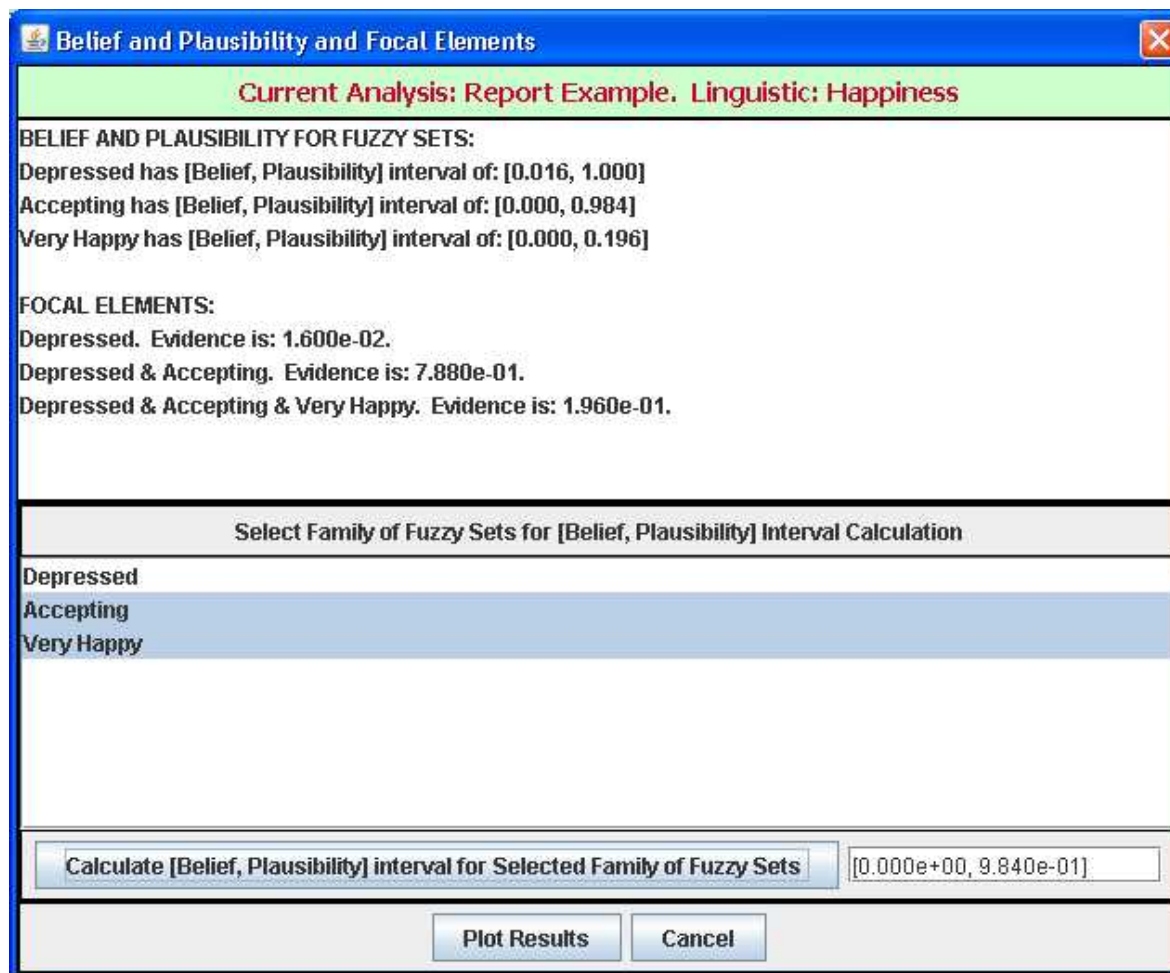
Good
Not so Good

Calculate [Belief, Plausibility] interval for Selected Family of Fuzzy Sets

Plot Results Cancel

Combination of Linguistic Variables: Example

Evaluate Variable Happiness for John



Belief and Plausibility and Focal Elements

Current Analysis: Report Example. Linguistic: Happiness

BELIEF AND PLAUSIBILITY FOR FUZZY SETS:
Depressed has [Belief, Plausibility] interval of: [0.016, 1.000]
Accepting has [Belief, Plausibility] interval of: [0.000, 0.984]
Very Happy has [Belief, Plausibility] interval of: [0.000, 0.196]

FOCAL ELEMENTS:
Depressed. Evidence is: 1.600e-02.
Depressed & Accepting. Evidence is: 7.880e-01.
Depressed & Accepting & Very Happy. Evidence is: 1.960e-01.

Select Family of Fuzzy Sets for [Belief, Plausibility] Interval Calculation

Depressed
Accepting
Very Happy

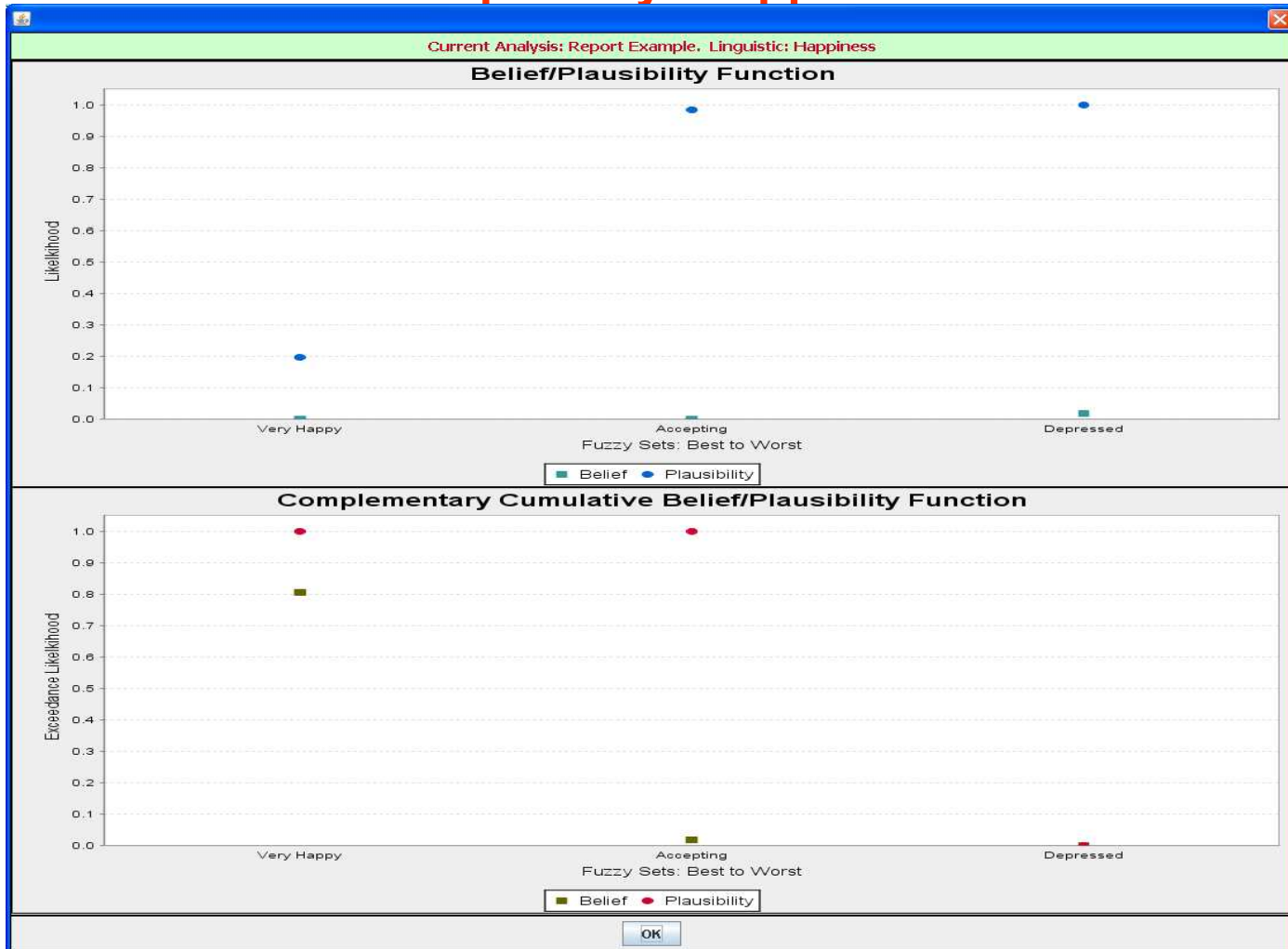
Calculate [Belief, Plausibility] interval for Selected Family of Fuzzy Sets [0.000e+00, 9.840e-01]

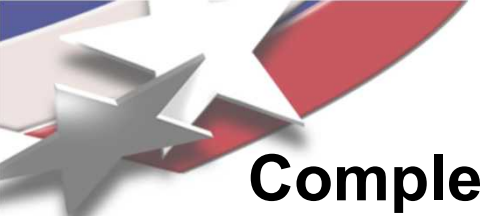
Plot Results **Cancel**

jldarby@sandia.gov

Combination of Linguistic Variables: Example

Summarize Results Graphically: Happiness for “John”



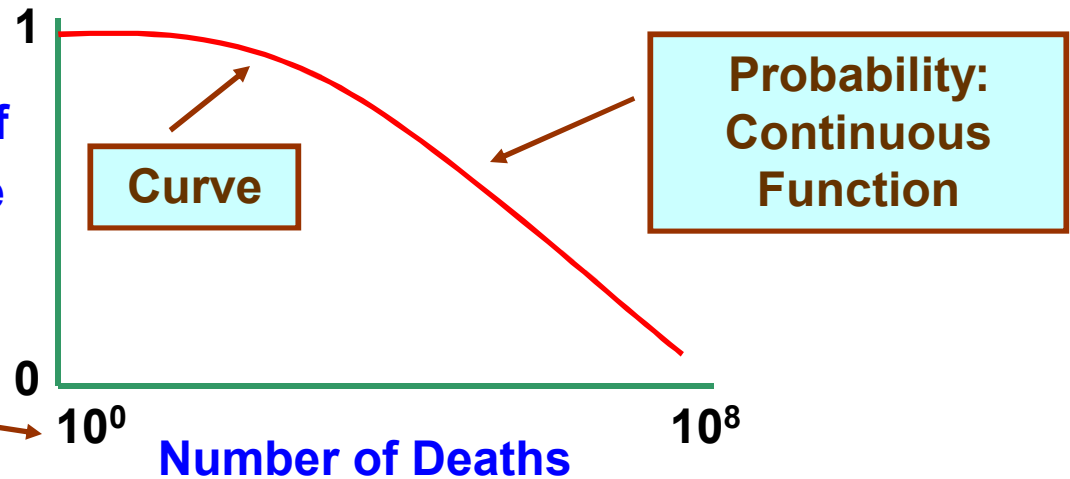


Results Expressed as Complementary Cumulative Belief/Plausibility Functions (CCBPFs) for Linguistic Variable

- Linguistic Fuzzy Sets **Ordered** from “Best” to “Worst”
 - CCBPFs are **Non-Increasing**
- “Likelihood” of **Exceeding** Fuzzy Set
- “Likelihood” is **Belief/Plausibility Interval**
- Analogous to Complementary Cumulative Probability Function (CCPF) for Random Variable
 - CCPF Random Variable is a real number
 - discrete or continuous
 - CCBPFs Variable has linguistic fuzzy sets
 - Discrete
 - CCPF is a One Function: a Curve
 - CCBPFs are Two Functions: an Interval

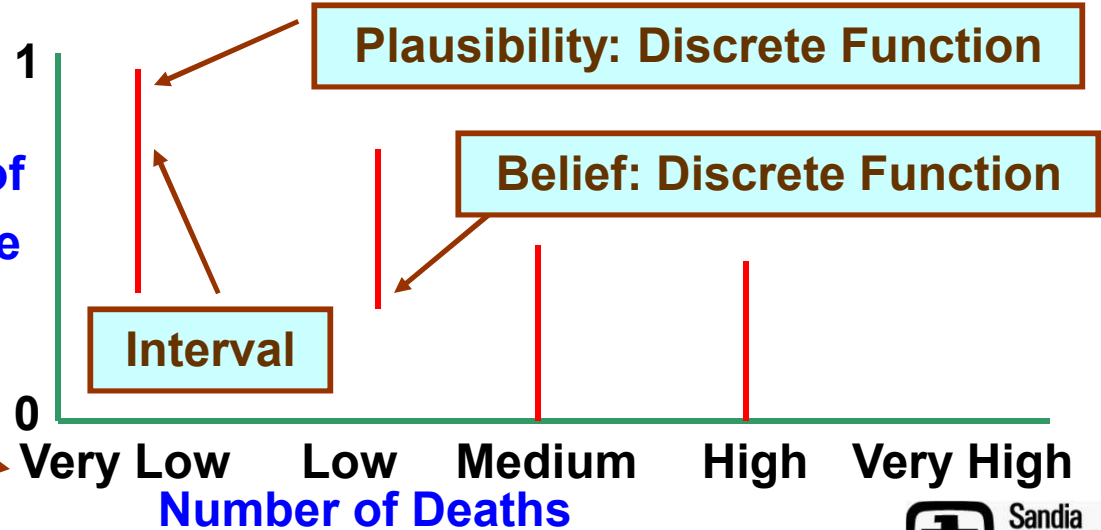
CCPF and CCBPFs

- **CCPF** Likelihood of Exceedance



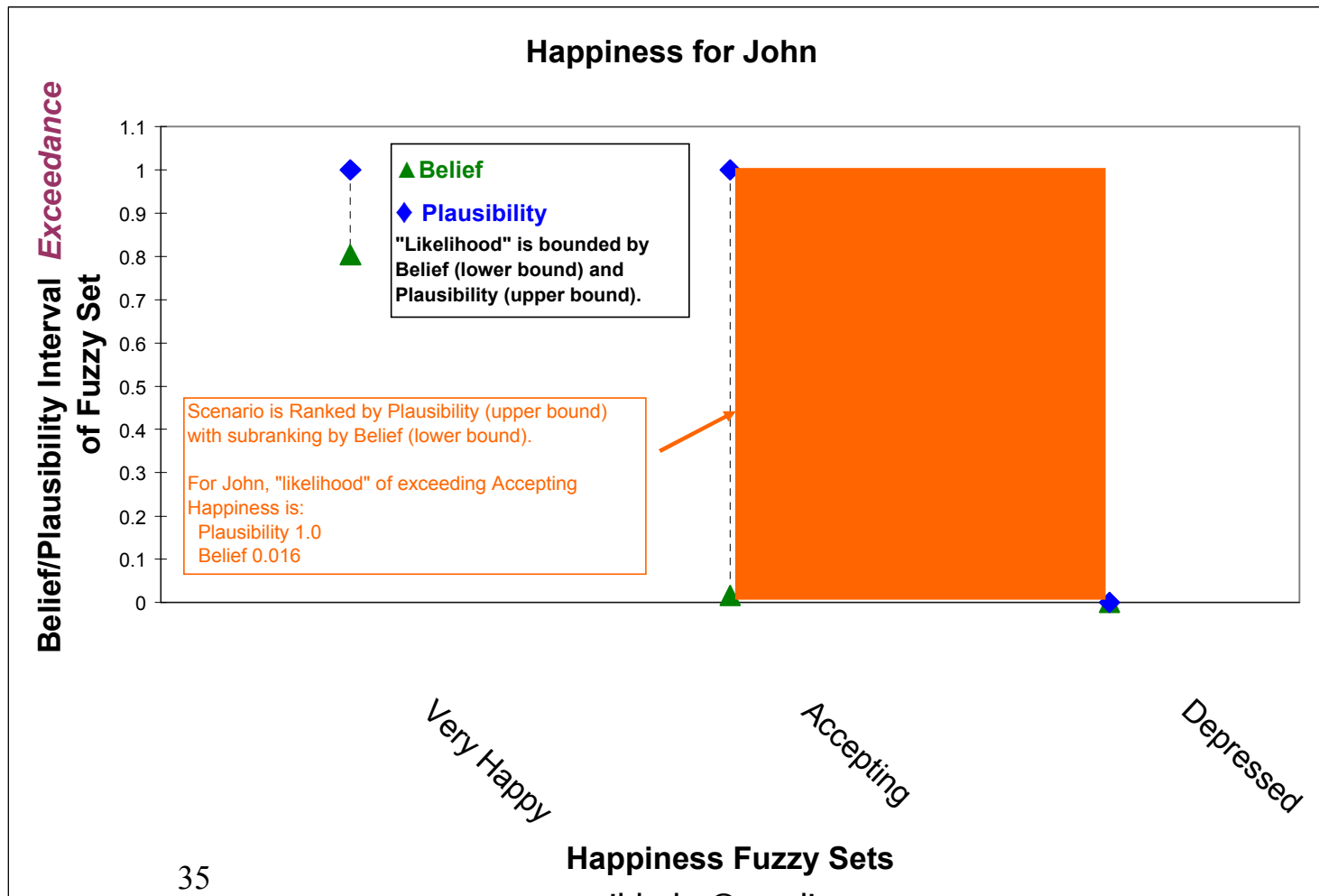
- **CCBPFs**

Likelihood of Exceedance



Combination of Linguistic Variables: Example

Graphical Summary for Ranking: "John"





Custom Software Tools: Java

- **BeliefConvolution**
 - Convolution of **Numeric** Variables with Belief/Plausibility
- **LinguisticBelief**
 - Evaluation of **Linguistic** Variables
 - Linguistic Fuzzy Sets
 - Approximate Reasoning
 - Belief/Plausibility
- **PoolEvidence**
 - Multiple Experts provide Evidence for Variables
 - Linguistic Fuzzy Sets
 - Combine Evidence
 - Pooled Evidence for Variables
 - Input for LinguisticBelief

LinguisticBelief: Example Application

LinguisticBelief Application

File Utilities Help

New Open Save EXIT! Exit

Current Analysis: National Planning Scenario

- Rule Linguistics
 - Long Term Consequence
 - Casualties (Equivalent Dead)
 - Adversary Estimate Consequence
 - Adversary Estimate Information Required
 - Detect Adversary Gathering Information
 - Adversary Estimate Scientific/Engineering Attributes Required
 - Detect Adversary Gathering Scientific/Engineering Attributes
 - Adversary Estimate Attributes Required
 - Adversary Estimate Resources Required
 - Detect Adversary Gathering Attributes
 - Immediate Consequence
 - Detect Adversary Gathering Resources
 - Vulnerability
 - Consequence
 - Adversary Estimate Vulnerability
 - Threat
 - Risk**
- Basic Linguistics
 - Dead
 - Injured/Ill
 - Infrastructure Damage
 - Evacuation
 - Contamination
 - National Economic Impact

Information for Selected Tree Node

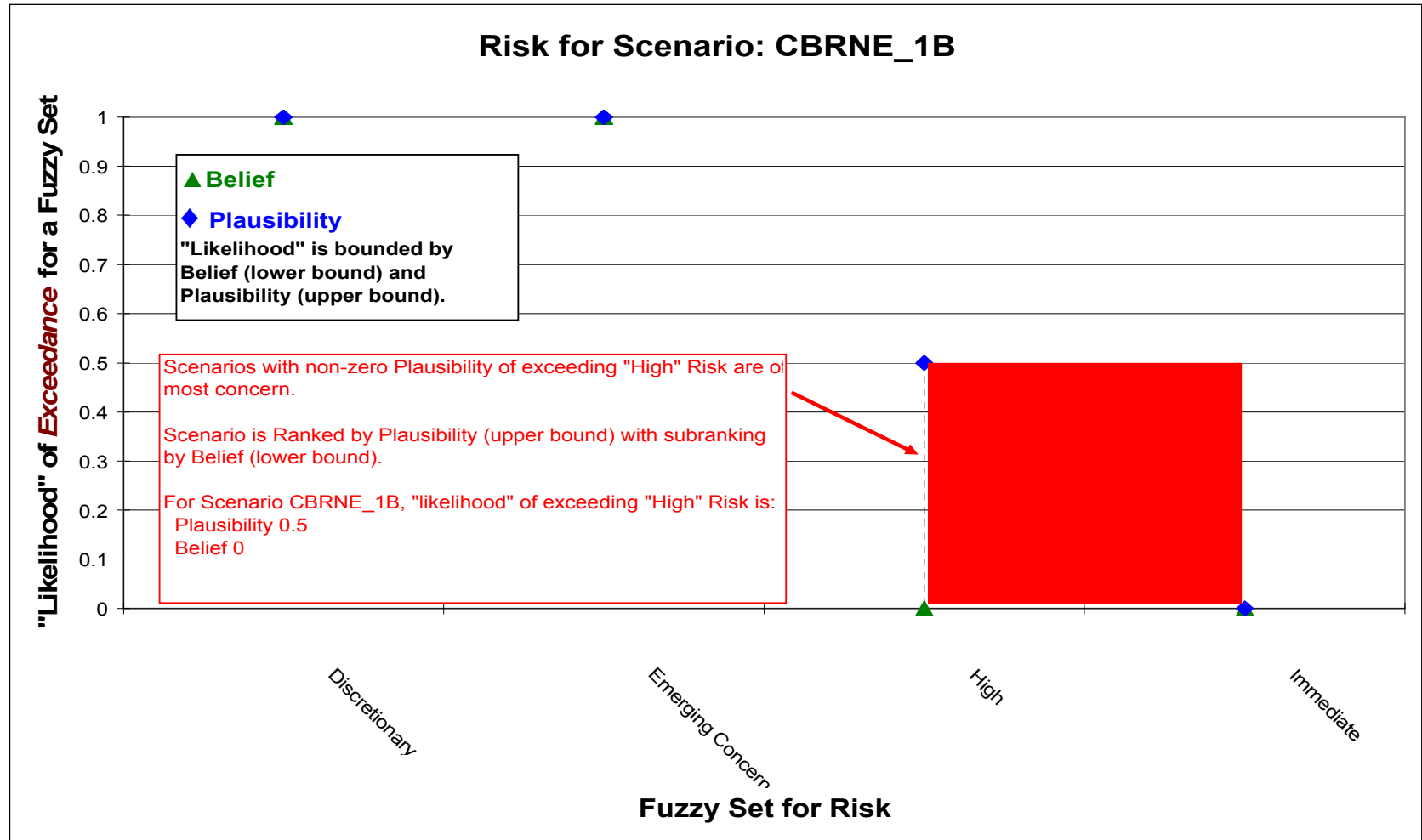
fuzzy sets: credible, Very Low, medium, -- Rule output fuzzy set: Negligible*, * Rule input fuzzy sets: credible, Very Low, high, -- Rule output fuzzy set: Negligible*, * Rule input fuzzy sets: credible, Very Low, very high, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, Low, very low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, Low, low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, Low, medium, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, Low, high, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, Low, very high, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, Medium, very low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, Medium, low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, Medium, medium, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, Medium, high, -- Rule output fuzzy set: Medium*, * Rule input fuzzy sets: credible, Medium, very high, -- Rule output fuzzy set: High*, * Rule input fuzzy sets: credible, High, very low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, High, low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: credible, High, medium, -- Rule output fuzzy set: Medium*, * Rule input fuzzy sets: credible, High, high, -- Rule output fuzzy set: High*, * Rule input fuzzy sets: credible, High, very high, -- Rule output fuzzy set: Extreme*, * Rule input fuzzy sets: likely, Very Low, very low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, Very Low, medium, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, Very Low, high, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, Very Low, very high, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, Low, very low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, Low, low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, Low, medium, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, Low, high, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, Low, very high, -- Rule output fuzzy set: Medium*, * Rule input fuzzy sets: likely, Medium, very low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, Medium, low, -- Rule output fuzzy set: Medium*, * Rule input fuzzy sets: likely, Medium, medium, -- Rule output fuzzy set: High*, * Rule input fuzzy sets: likely, Medium, high, -- Rule output fuzzy set: High*, * Rule input fuzzy sets: likely, Medium, very high, -- Rule output fuzzy set: High*, * Rule input fuzzy sets: likely, High, very low, -- Rule output fuzzy set: Low*, * Rule input fuzzy sets: likely, High, low, -- Rule output fuzzy set: Medium*, * Rule input fuzzy sets: likely, High, medium, -- Rule output fuzzy set: High*, * Rule input fuzzy sets: likely, High, high, -- Rule output fuzzy set: Extreme*, * Rule input fuzzy sets: likely, High, very high, -- Rule output fuzzy set: Extreme*, . [Belief, Plausibility] Intervals for fuzzy sets are as follows : Fuzzy Set: Negligible [0.000, 0.384], Fuzzy Set: Low [0.000, 0.349], Fuzzy Set: Medium [0.155, 1.000], Fuzzy Set: High [0.000, 0.720], Fuzzy Set: Extreme [0.000, 0.000], .



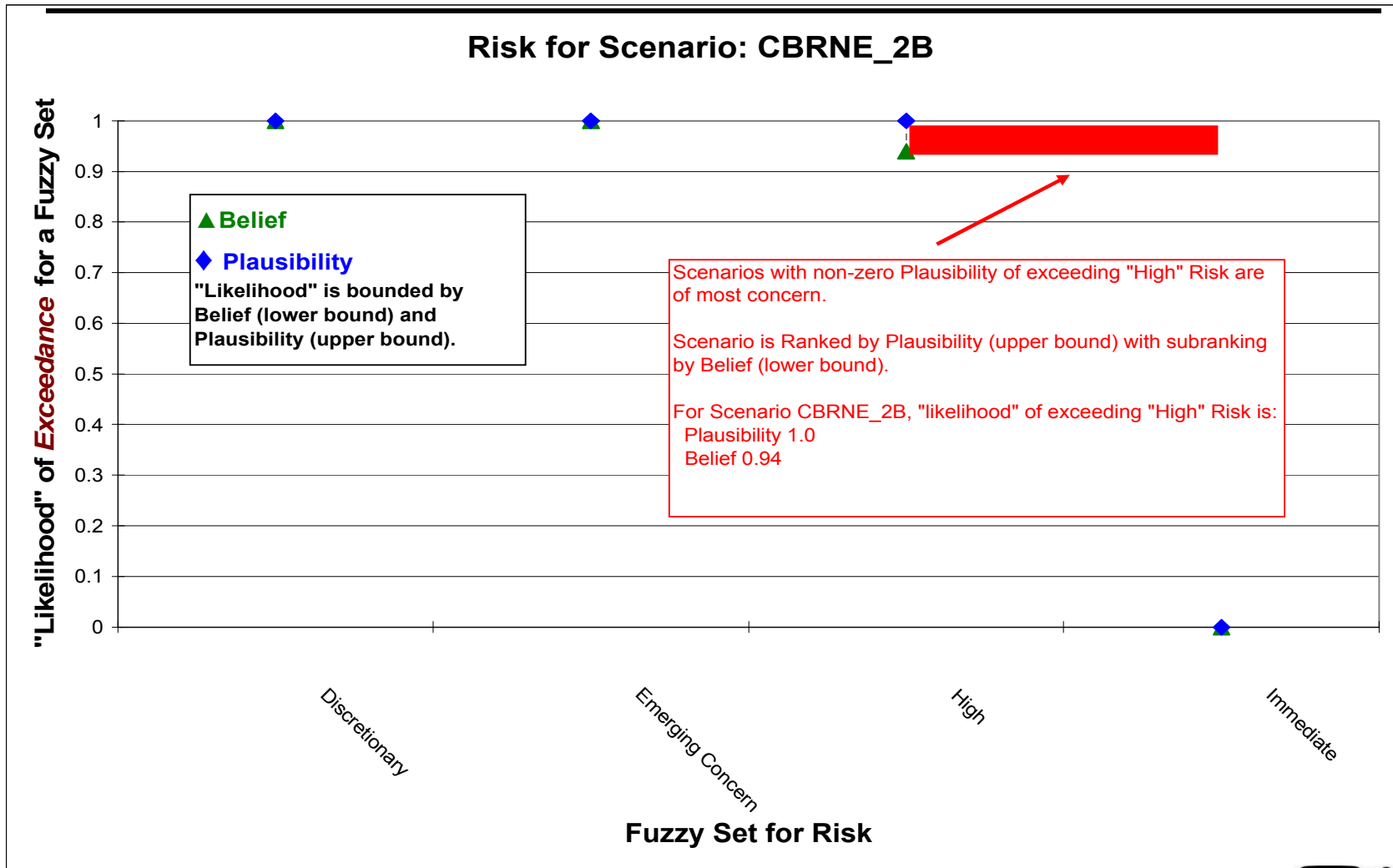
Rank Order Scenarios by Risk

- **Rank a scenario by the Highest Non-Zero Plausibility of Exceeding the “Worst” Fuzzy Set**
 - **For Scenarios with Equal Plausibility, Subrank by Highest Belief**
- **Extension of “Probability of Exceedance” approach**
 - **Uses Fuzzy Sets instead of Numbers**
 - **Uses Belief/Plausibility Interval instead of Probability**
- **Can be “Color Coded”**
 - **Shown for 3 of 5 scenarios in Following from SAND2007-1301**

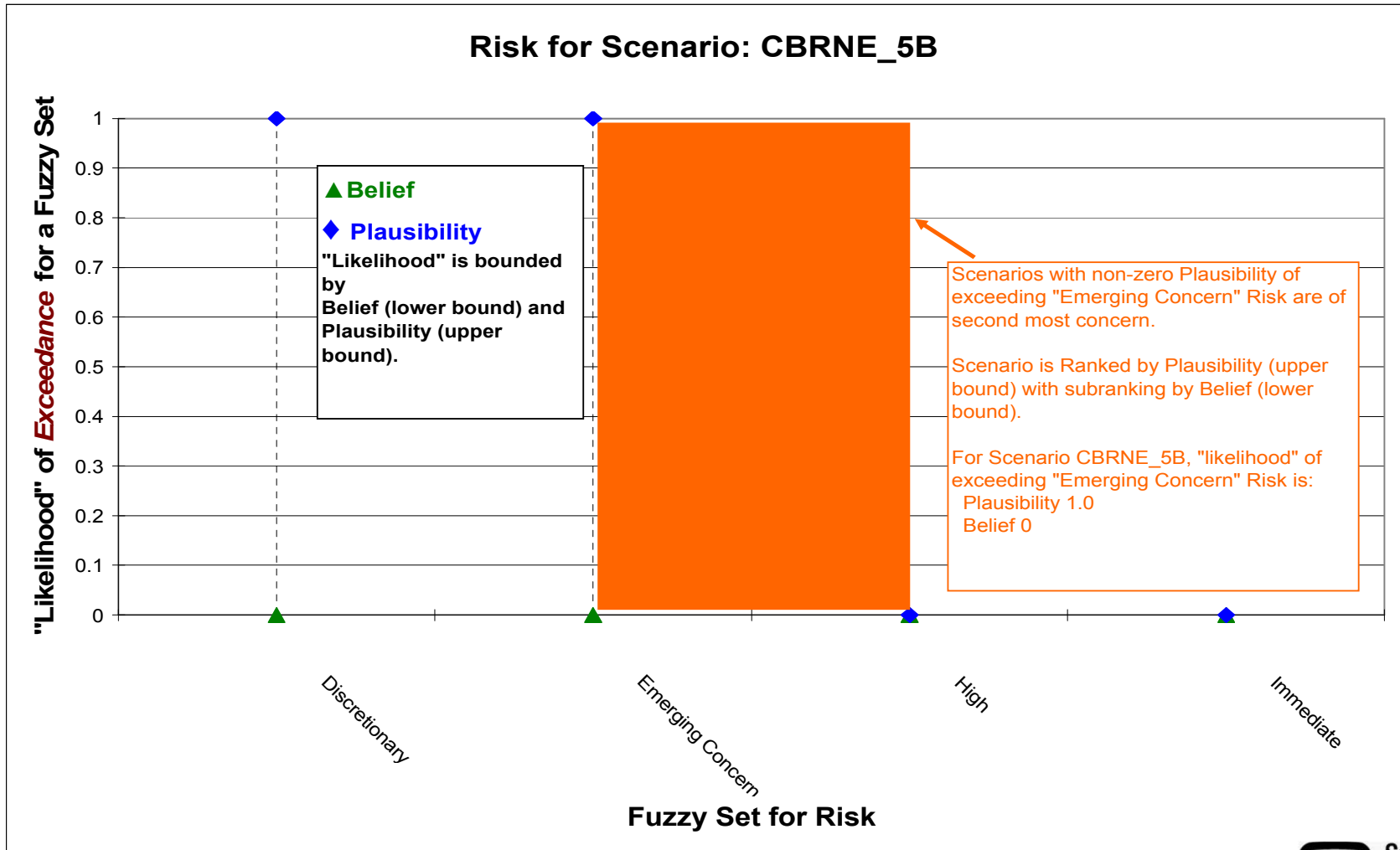
Rank Order Results



Rank Order Results



Rank Order Results





Results of Ranking of All Five Scenarios

RANKING FOR SCENARIOS CBRNE_1B through CBRNE_5B

For **Exceeding** Fuzzy Set “**High**” the Scenarios rank ordered (decreasing) are:

CBRNE_2B has plausibility of exceedance of 1.0 and belief of exceedance of 0.94

CBRNE_3B has plausibility of exceedance of 1.0 and belief of exceedance of 0.77

CBRNE_4B has plausibility of exceedance of 1.0 and belief of exceedance of 0.64

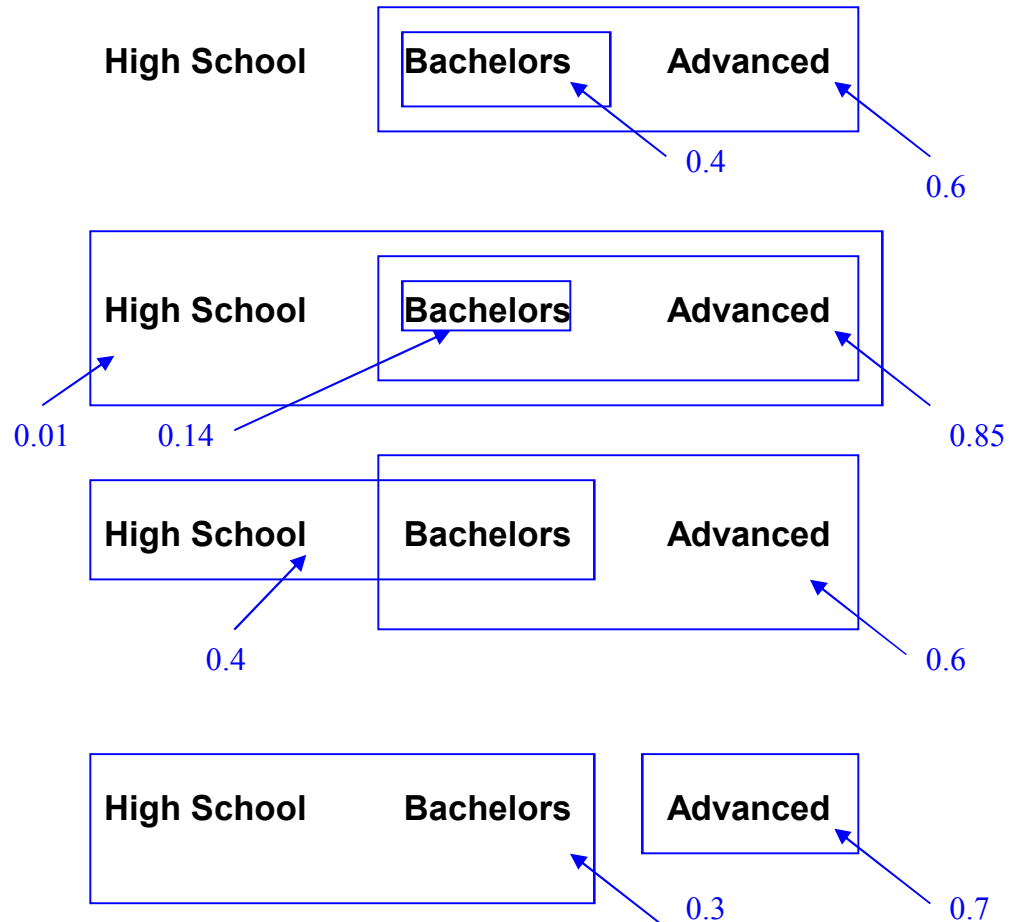
CBRNE_1B has plausibility of exceedance of 0.5 and belief of exceedance of 0.0

For **Exceeding** Fuzzy Set “**Emerging Concern**” the Scenarios rank ordered (decreasing) (not already ranked for a worse fuzzy set) are:

CBRNE_5B has plausibility of exceedance of 1.0 and belief of exceedance of 0.0

Pool Evidence from Many Experts

**Adversary Level of
Technical Training:
4 Experts Assign
Evidence**



PoolEvidence: Example Application

Pooled Evidence Application

File Help

New Open Save EXIT! Exit

Current Analysis: None

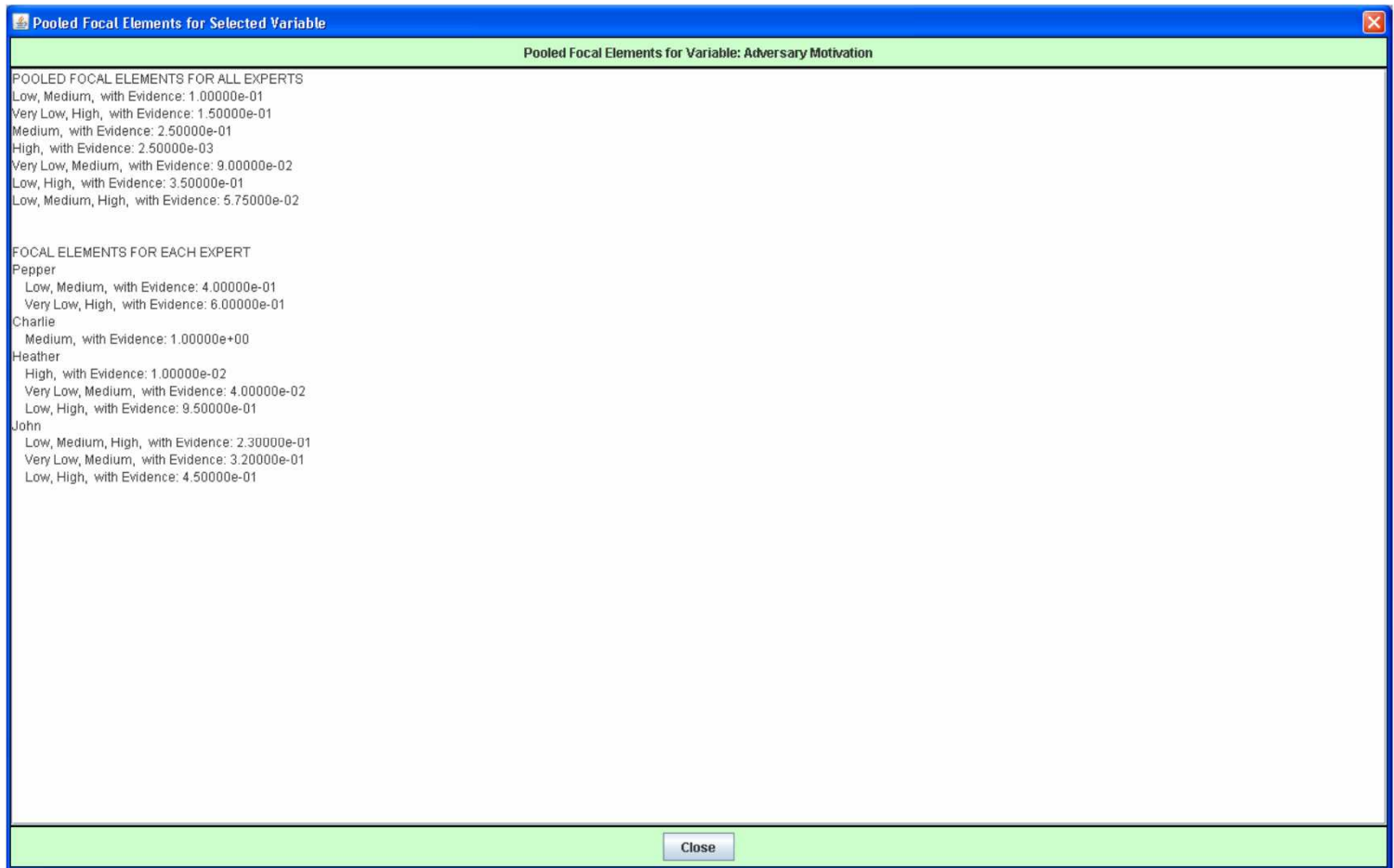
| Variables for Analysis | Fuzzy Sets for Selected Variable |
|---|---|
| <p>Variables for Current Analysis: Actual Test</p> <ul style="list-style-type: none">Adversary MotivationDefender ResourcesResult | <p>Fuzzy Sets for Variable: Adversary Motivation</p> <ul style="list-style-type: none">Very LowLowMediumHighVery High |

Show Pooled Focal Elements for Selected Variable

| Experts for Selected Variable | Focal Elements for Selected Expert |
|---|---|
| <p>Experts for Variable: Adversary Motivation</p> <ul style="list-style-type: none">PepperCharlieHeatherJohn | <p>Focal Elements for Expert: John</p> <ul style="list-style-type: none">Low, Medium, High, with Evidence: 2.30000e-01Very Low, Medium, with Evidence: 3.20000e-01Low, High, with Evidence: 4.50000e-01 |

Close

PoolEvidence





Summary

- **Uncertainty**
 - **Belief/Plausibility for Epistemic Uncertainty**
- **Qualitative Variables**
 - **Purely Linguistic Fuzzy Sets**
- **Combining Variables**
 - **Approximate Reasoning**
 - **Convolution with Belief/Plausibility**
- **Custom Software Tools**
 - **BeliefConvolution**
 - **LinguisticBelief**
 - **PoolEvidence**