When There Is Little Data Can AI Still Work?
-- Approximate Knowledge Interpolation and its Applications

Professor Qiang Shen FREng
Department of Computer Science
Aberystwyth University, UK

Outline

- Introduction
- Reasoning with limited (imprecise) data
  - Underlying approach: T-FRI
- Extended approaches:
  - Adaptive T-FRI
  - Backward T-FRI
  - Higher-order T-FRI
  - Dynamic T-FRI
  - Weighted T-FRI
- Example applications:
  - Computer network security
  - Mammogram mass analysis
- Conclusion

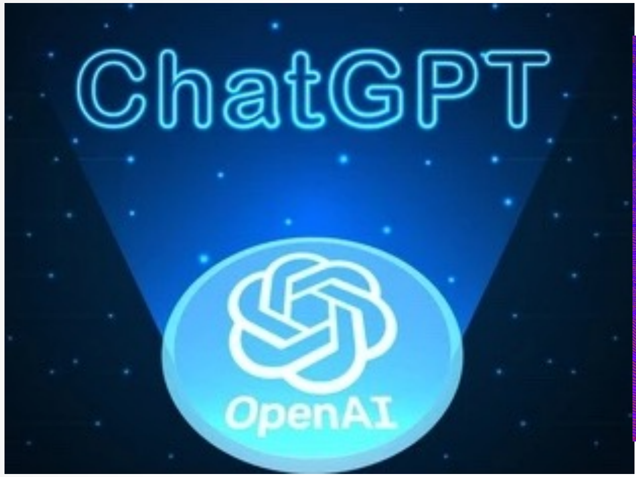
T-FRI: Transformation-based fuzzy rule interpolation

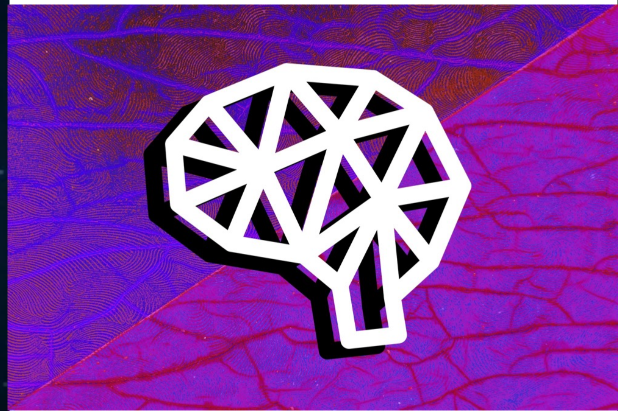
Potential of AI

- Estimated $15.7 trillion impact on global economy by 2030
- Performing tasks that once thought only humans could do
    - Simple forms: scheduling meetings, ordering meals, answering questions about the weather
    - Advanced forms: building self-driving cars, diagnosing medical conditions, bringing kids to schools
- Top trends for AI and AI-related development
    - Natural language communication between human and machines
    - Strengthening and disruption of cybersecurity
    - Integration of AI algorithms with quantum computation
    - Governance on AI ethical principles and ethical data collection
    - Diversification of AI models to build culturally sensitive AI systems
    - Demystification of complex algorithms: Explainable AI

Where Powerful AI May Collapse

- AI, especially deep learning is effective when working



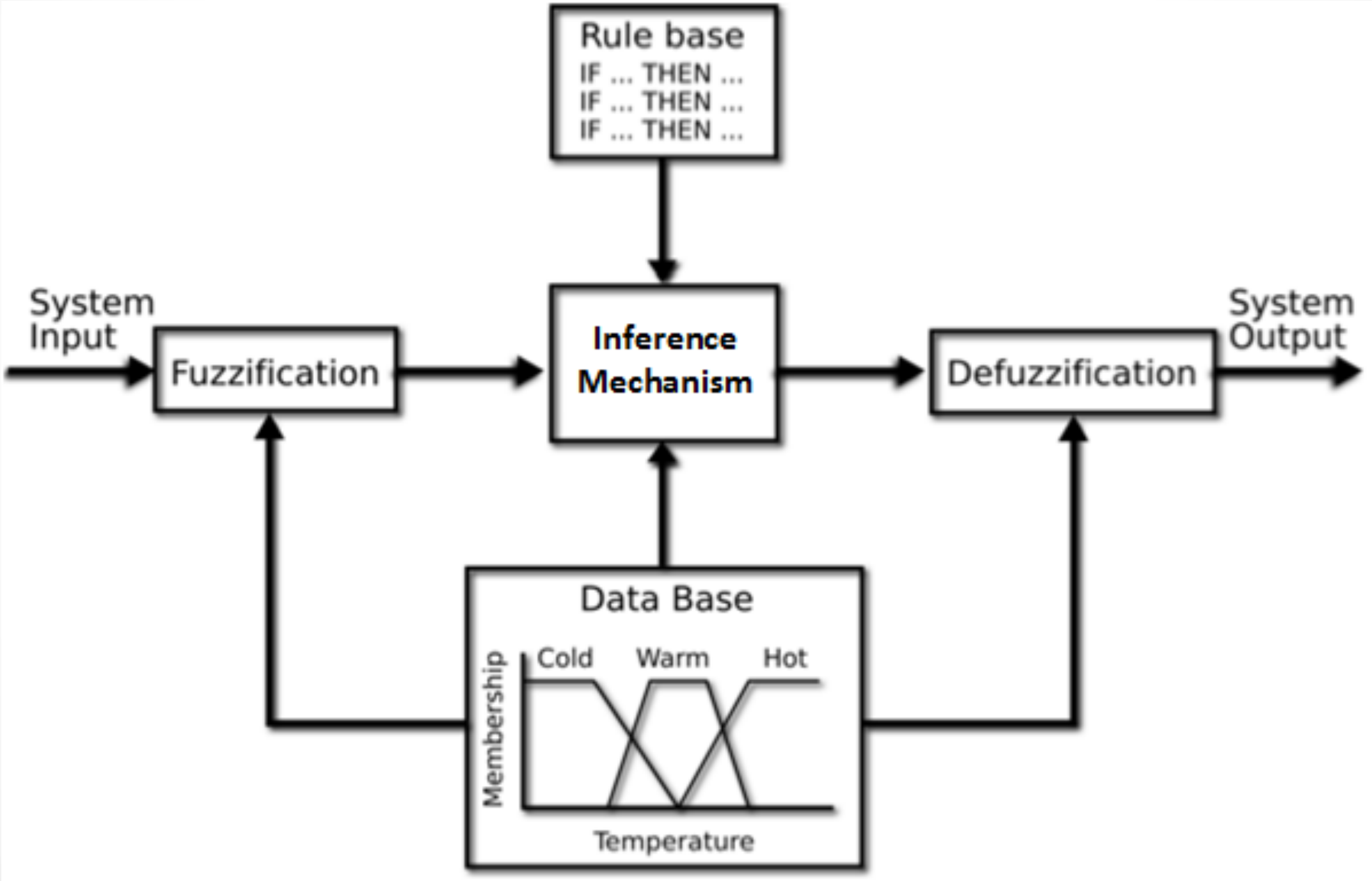


SEER

DALL-E2

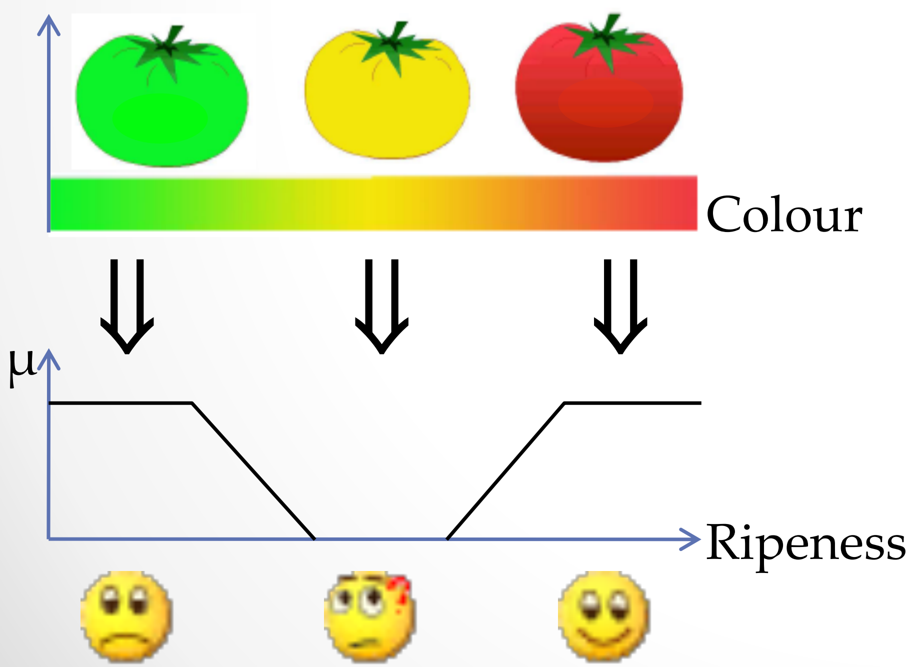
- Deep learning becomes difficult when required to explain its outcomes
  - For imprecisely described problems, in particular.
- AI stops working when facing novel problems with little training data
  - For generative AI and big data driven techniques.

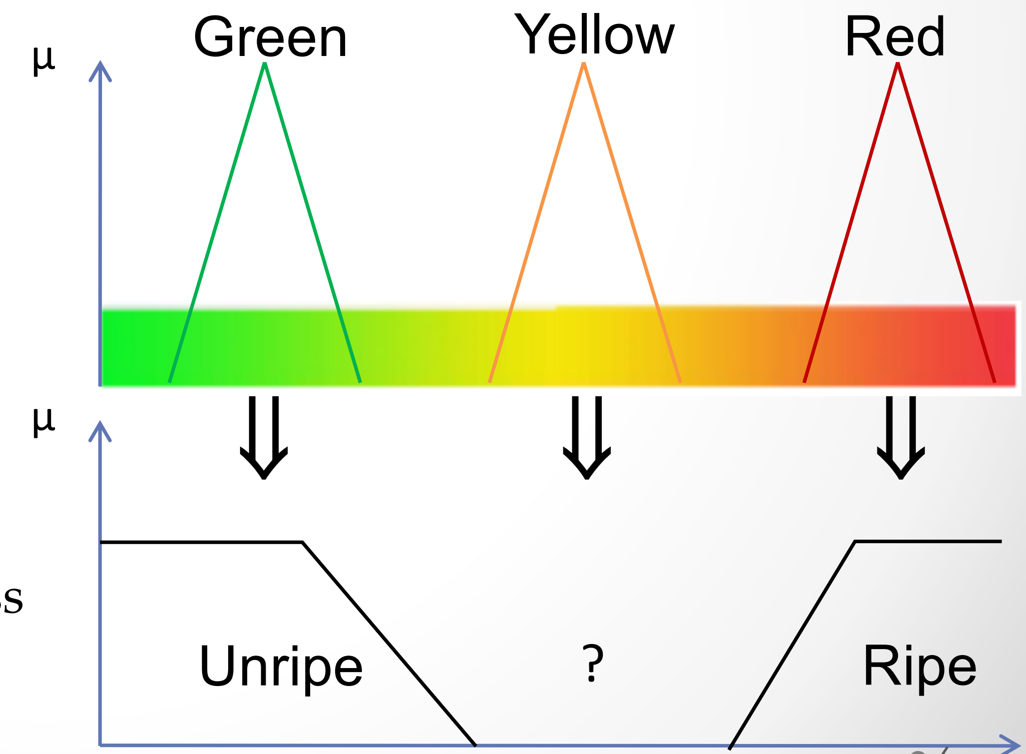
Fuzzy Inference Systems
(Dense Rule Bases)



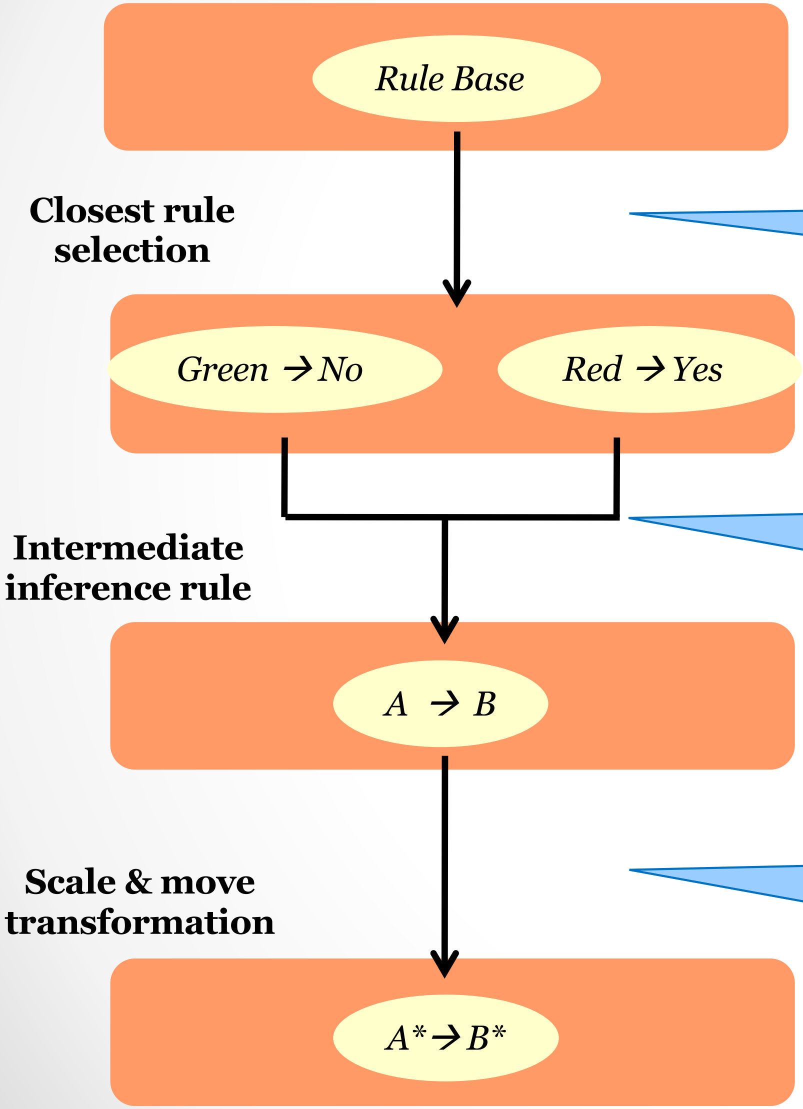
Fuzzy Rule Interpolation
(Sparse Rule Bases)

- If a given observation does not (partially) match any rule, conventional fuzzy inference does not work
- Classical "tomato classification" problem:





Transformation-Based Fuzzy Rule Interpolation (T-FRI)

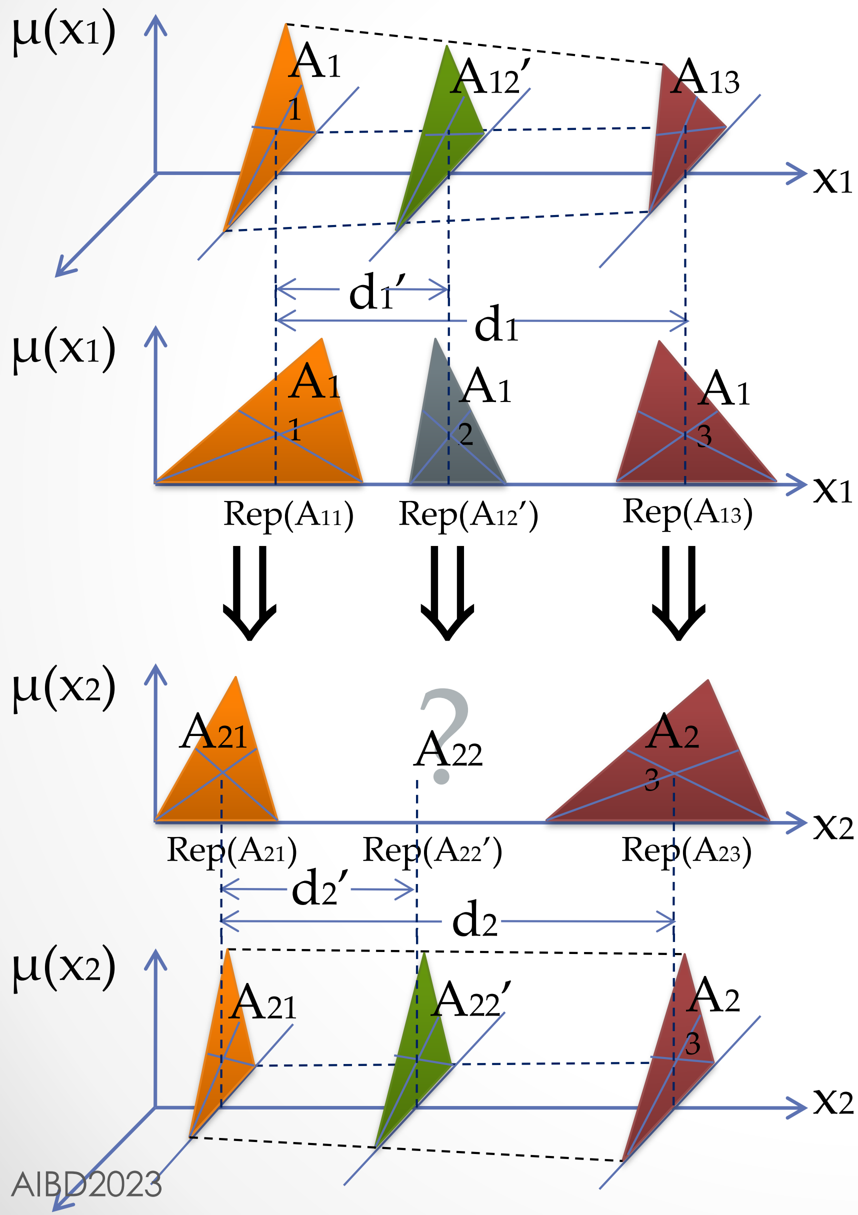


Select nearest neighboring rules, w.r.t. a given unmatched observation A*

Construct a new intermediate rule by linearly interpolating selected neighboring rules, guided by A*

Derive conclusion B* via scale and move transformations, ensuring similarity

Construction of Intermediate Rule



Representative value:

$\mathrm{Rep}(A_{ij}) =$ Centre of gravity

Relative placement factor:

$$\lambda_1 = \frac{d_1'}{d_1} = \frac{d_2'}{d_2}$$

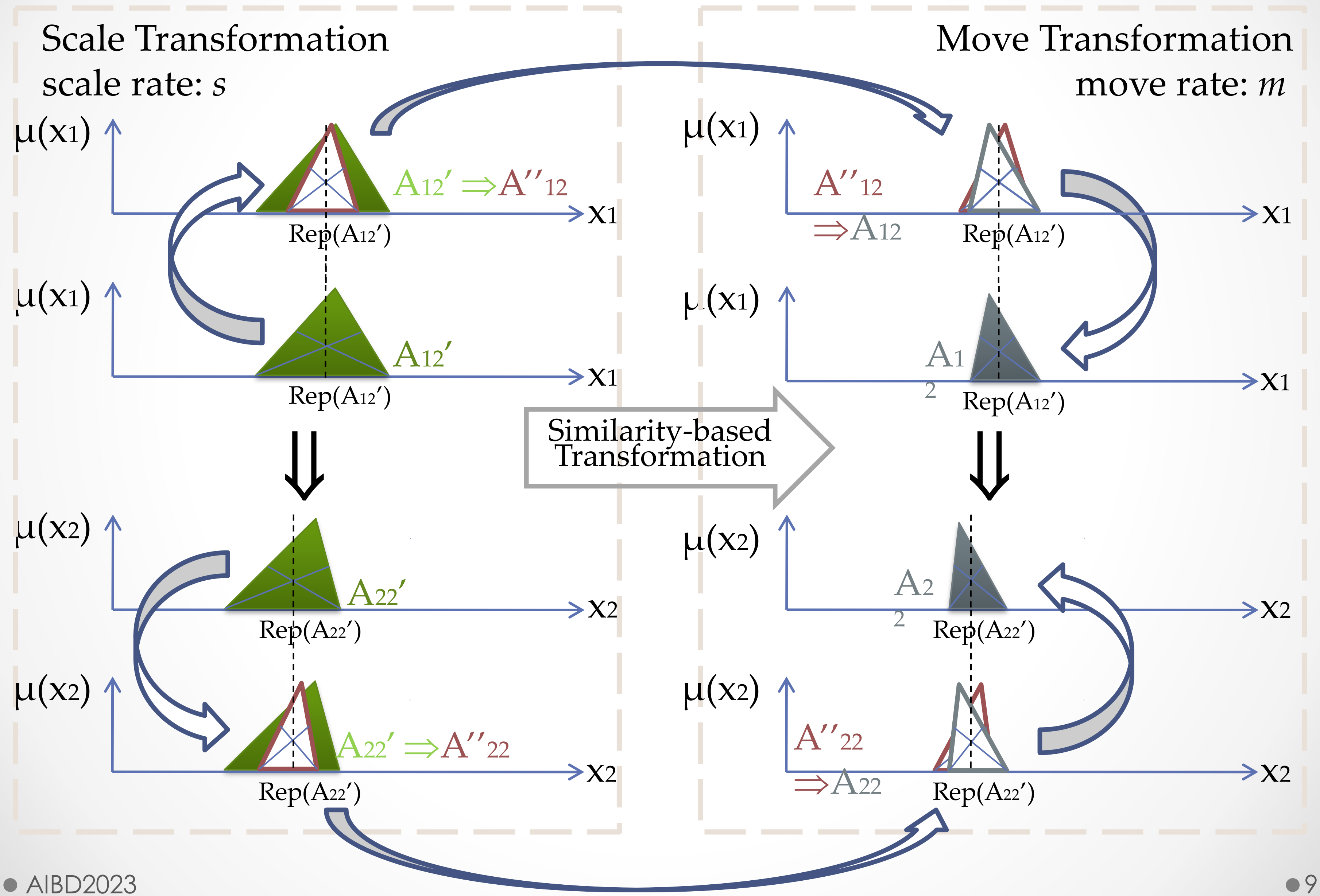
⇔

Rule: If x_1 is A'_{12}, then x_2 is A'_{22}

O: x_1 is A_{12}

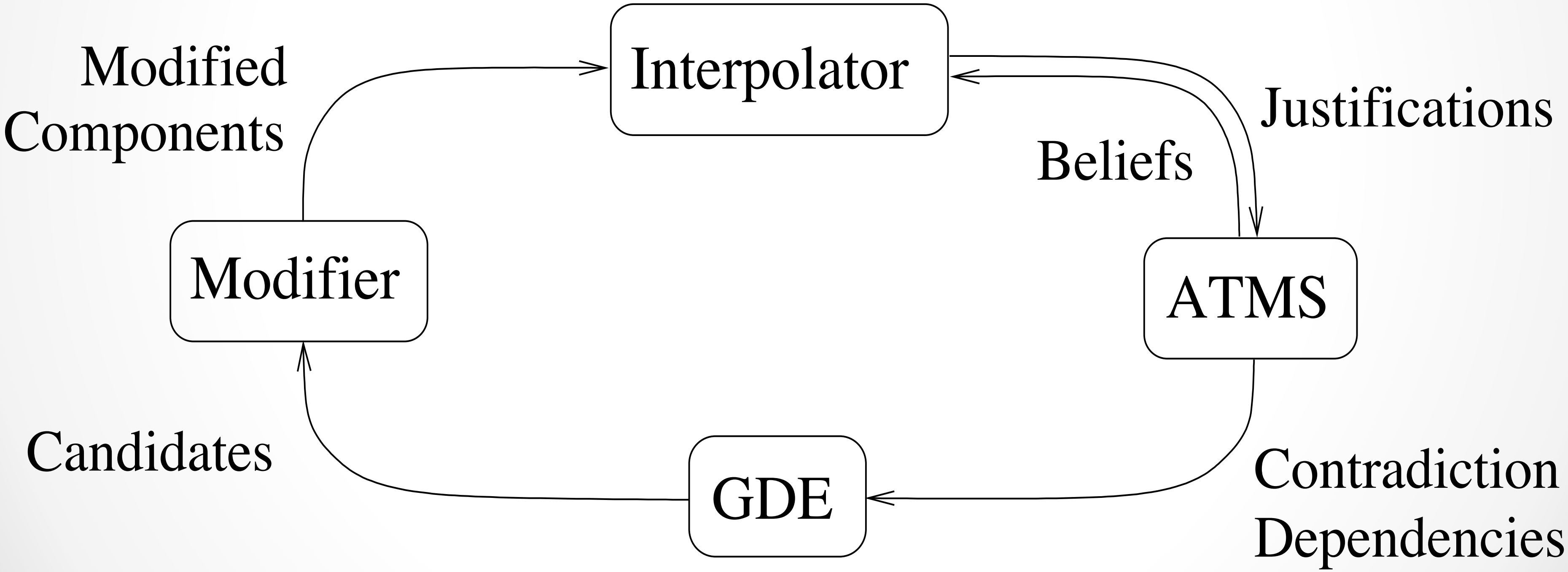
C: x_2 is A_{22}

Scale and Move Transformations



Adaptive T-FRI

- Problem: Inconsistent interpolated results due to rule sparseness and error propagation
- Solution: Rectification using symbolic AI: ATMS and GDE

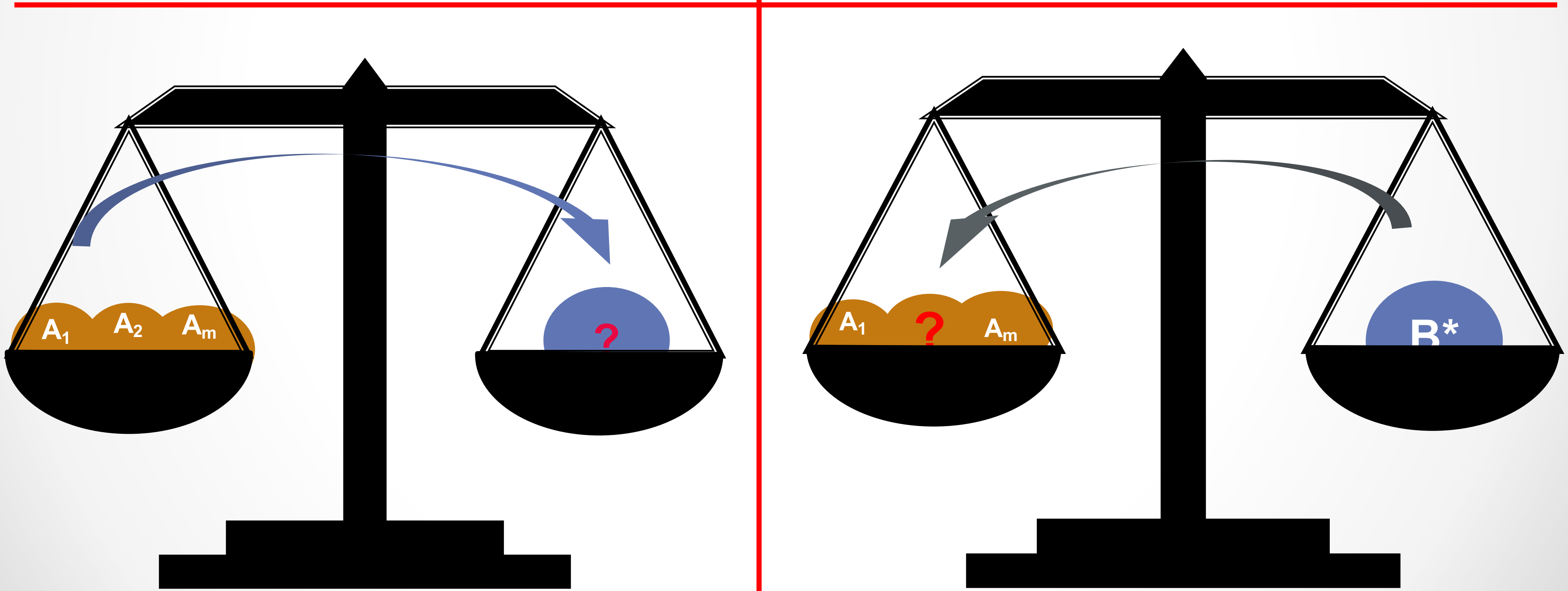


ATMS: Assumption-based truth maintenance system
GDE: General diagnostic engine

Backward T-FRI

- Problem: Missing antecedent values in observations
- Solution: Reverse inference using known antecedent and consequent values

Forward-FRI | **Backward-FRI**

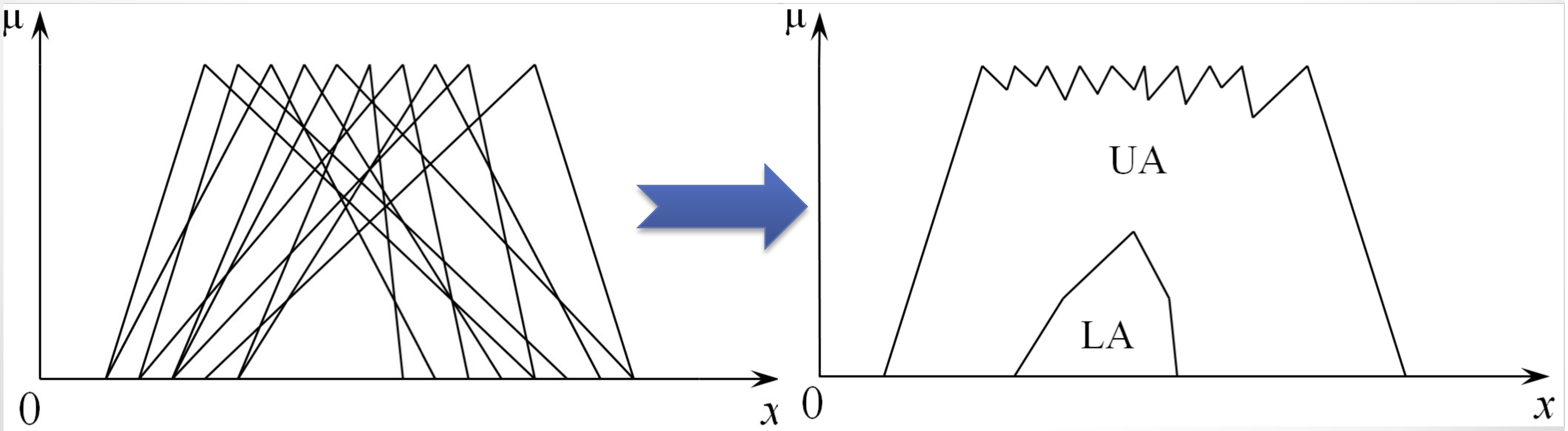


$$B_l^* = f_{T-FRI}((A_1^*,\cdots,A_l^*,\cdots,A_M^*),(R_i,\cdots,R_t))$$

$$A_l^* = f_{B-FRI}((B^*,A_1^*,\cdots,A_{l-1}^*,A_{l+1}^*,\cdots,A_M^*),(R_i,\cdots,R_t))$$

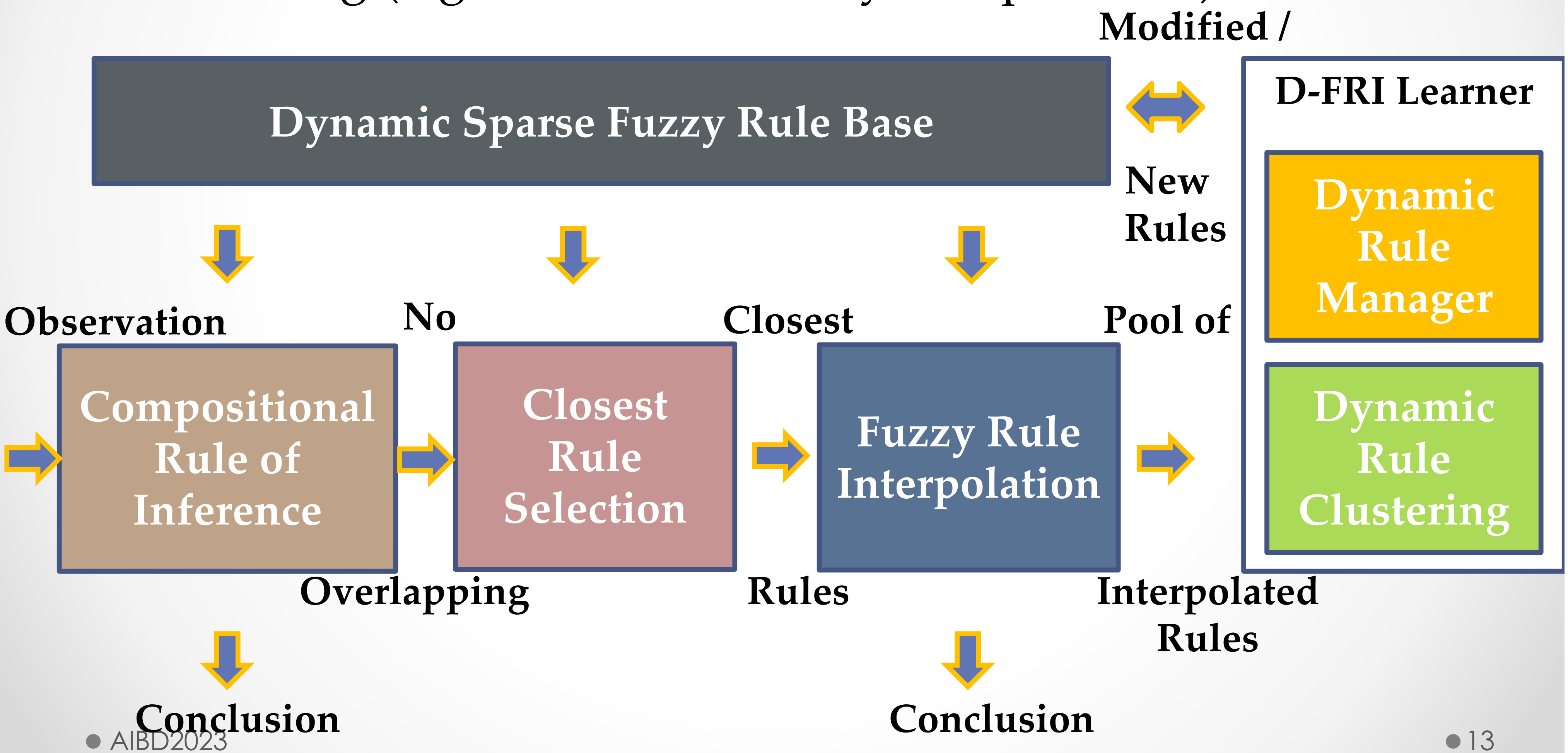
Higher-Order T-FRI

- Problem: Incapability of handling more than one form of uncertainty
- Solution: Harnessing additional uncertain information in rules/observations with rough-fuzzy sets



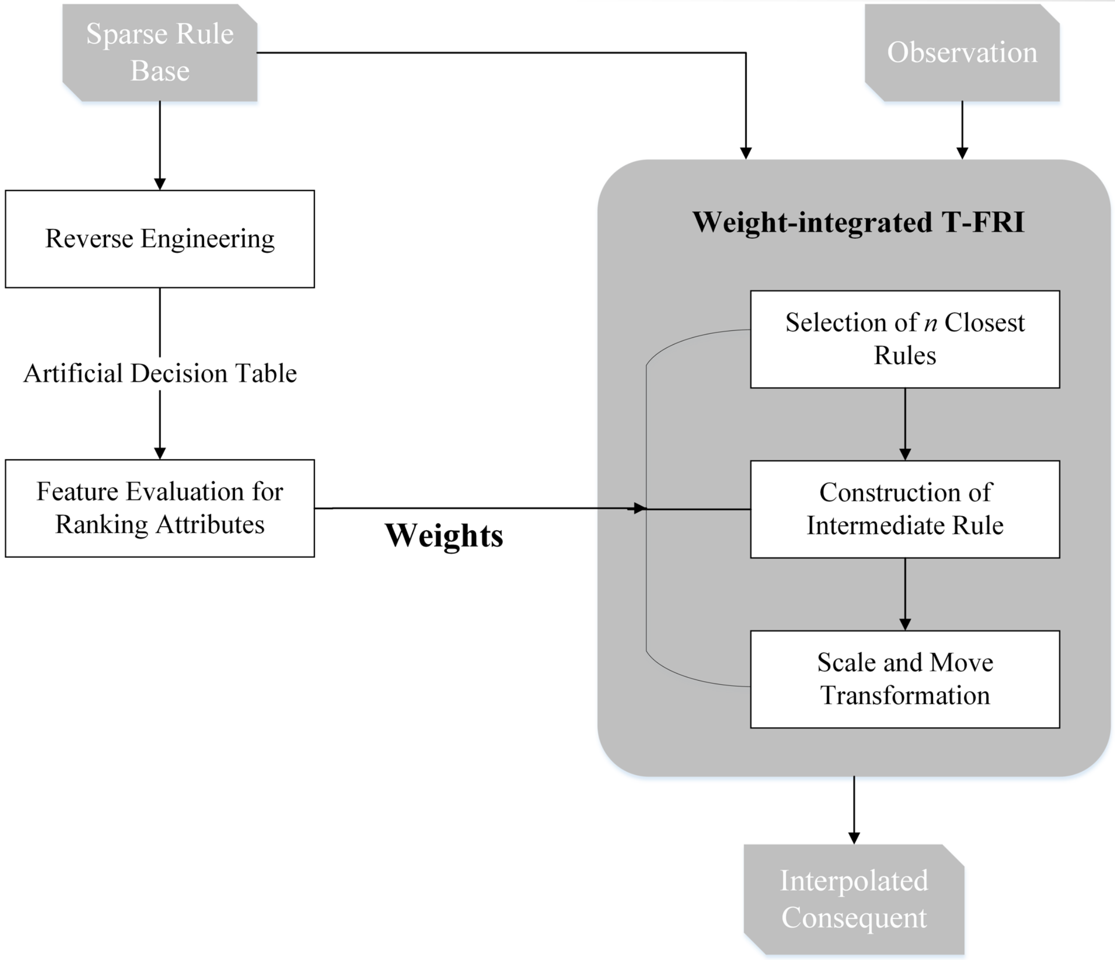
Dynamic T-FRI (D-FRI)

- Problem: Ineffective and/or inefficient interpolated results due to use of static rule base
- Solution: Learning from interpolated results with clustering (e.g., via evolutionary computation)



Weighted T-FRI (W-T-FRI)

- Problem: Unrealistic/ counter-intuitive knowledge representation with all antecedents assumed of equal significance
- Solution: Learning attribute weights from given sparse rules only via *reverse engineering*



Reverse Engineering: Turing Rul

- Identifying all possible antecedent variables and their value domains
- Expanding any rule with missing variables into a rule set s.t. each missing variable in every expanded rule takes one of its possible fuzzy values

Example sparse rule base:

1. If *Temperature* is *Hot* and *Outlook* is *Sunny*, then *Swimming*.
2. If *Temperature* is *Hot* and *Outlook* is *Cloudy*, then *Swimming*.
3. If *Outlook* is *Rain*, then *Weight lifting*.
4. If *Temperature* is *Mild* and *Wind* is *Windy*, then *Weight lifting*.
5. If *Temperature* is *Mild* and *Wind* is *Not Windy*, then *Volleyball*.

Ar

Tempe

H

H

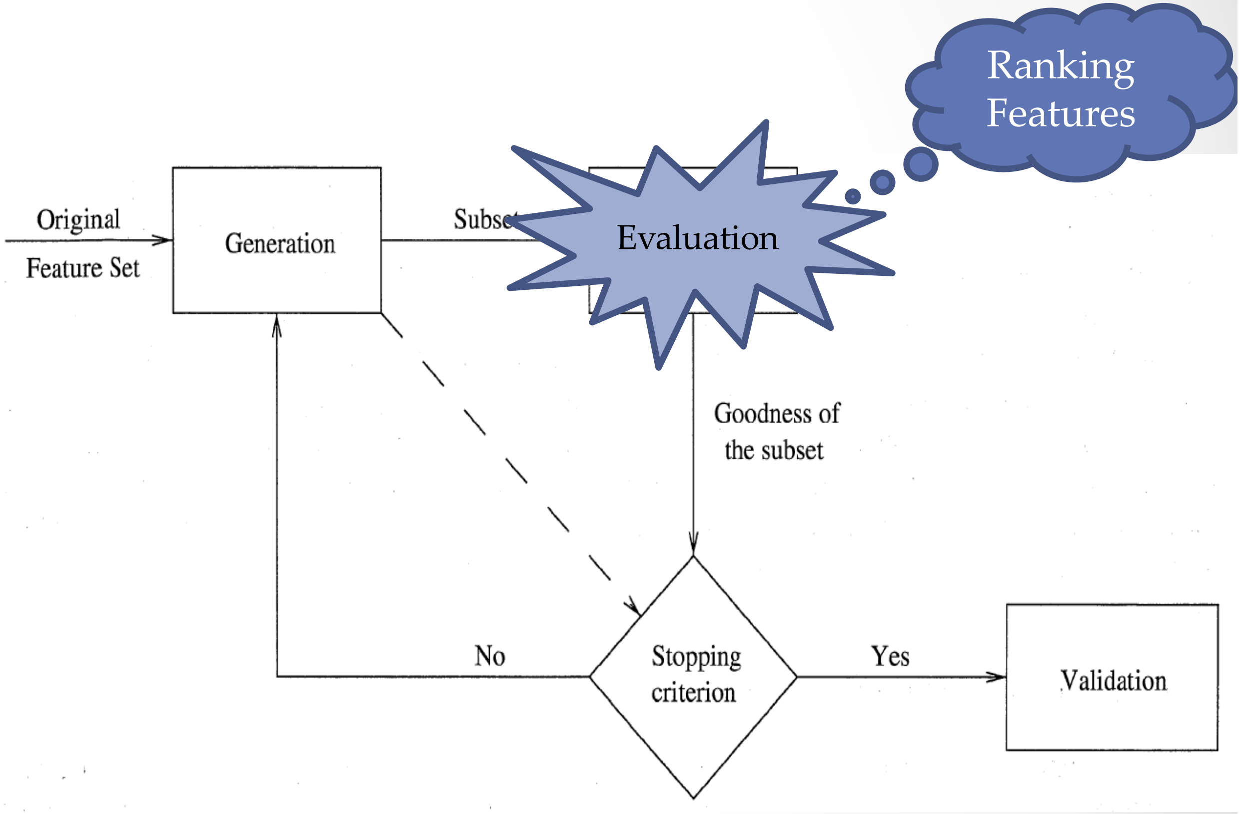
H

H

Temperature	Outlook	Humidity	Wind	Decision
Hot	Sunny	Humid	Windy	Swimming
Hot	Sunny	Humid	Not Windy	Swimming
Hot	Sunny	Normal	Windy	Swimming
Hot	Sunny	Normal	Not Windy	Swimming
Hot	Cloudy	Humid	Windy	Swimming
Hot	Cloudy	Humid	Not Windy	Swimming
Hot	Cloudy	Normal	Windy	Swimming
Hot	Cloudy	Normal	Not Windy	Swimming
Hot	Rain	Humid	Windy	Weight lifting
Hot	Rain	Humid	Not Windy	Weight lifting
Hot	Rain	Normal	Windy	Weight lifting
Hot	Rain	Normal	Not Windy	Weight lifting
Mild	Rain	Humid	Windy	Weight lifting
Mild	Rain	Humid	Not Windy	Weight lifting
Mild	Rain	Normal	Windy	Weight lifting
Mild	Rain	Normal	Not Windy	Weight lifting
Cool	Rain	Humid	Windy	Weight lifting
Cool	Rain	Humid	Not Windy	Weight lifting
Cool	Rain	Normal	Windy	Weight lifting
Cool	Rain	Normal	Not Windy	Weight lifting
Mild	Sunny	Humid	Windy	Weight lifting
Mild	Sunny	Normal	Windy	Weight lifting
Mild	Cloudy	Humid	Windy	Weight lifting
Mild	Cloudy	Normal	Windy	Weight lifting
Mild	Rain	Humid	Windy	Weight lifting
Mild	Rain	Normal	Windy	Weight lifting
Mild	Sunny	Humid	Not Windy	Volleyball
Mild	Sunny	Normal	Not Windy	Volleyball
Mild	Cloudy	Humid	Not Windy	Volleyball
Mild	Cloudy	Normal	Not Windy	Volleyball
Mild	Rain	Humid	Not Windy	Volleyball
Mild	Rain	Normal	Not Windy	Volleyball

Weight Generation via Feature Ranking

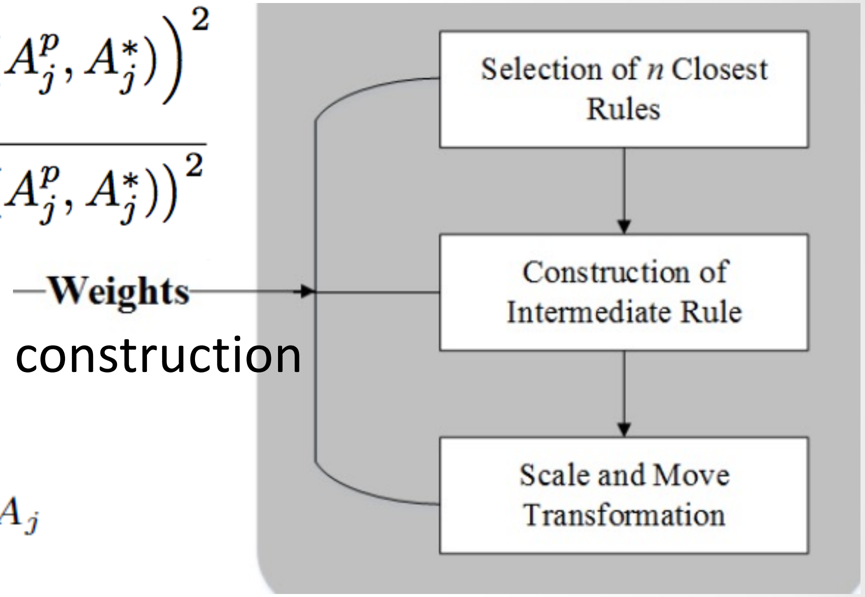
- Ranking through variable evaluation
- Evaluation methods often embedded in feature selection
- Feature selection methods:
  - Information gain-based
  - Relief-F metric-based
  - Laplacian score-based
  - Rough set-based
  - Correlation-based
  - Consistency-based
  - …



Implementation of W-T-FRI

- Weight-guided selection of closest rules

$$\tilde{d}(r^p, o^*) = \frac{1}{\sqrt{\sum_{t=1}^{m}\left(\frac{1-W_t}{m-1}\right)^2}}\sqrt{\sum_{j=1}^{m}\left(\left(\frac{1-W_j}{m-1}\right)d(A_j^p, A_j^*)\right)^2}$$
$$= \frac{1}{\sqrt{\sum_{t=1}^{m}(1-W_t)^2}}\sqrt{\sum_{j=1}^{m}\left((1-W_j)d(A_j^p, A_j^*)\right)^2}$$



- Weighted parameters for intermediate rule construction

$$\tilde{\hat{w}}_z^i = \sum_{j=1}^{m} W_j \hat{w}_j^i, \qquad \tilde{\delta}_z = \sum_{j=1}^{m} W_j \delta_{A_j}$$

- Weighted transformation

$$\tilde{s}_z = \sum_{j=1}^{m} W_j s_{A_j}, \qquad \tilde{m}_z = \sum_{j=1}^{m} W_j m_{A_j}$$

- Empirically, only two closest rules required in Weighted T-FRI

Application: Computer Network Security

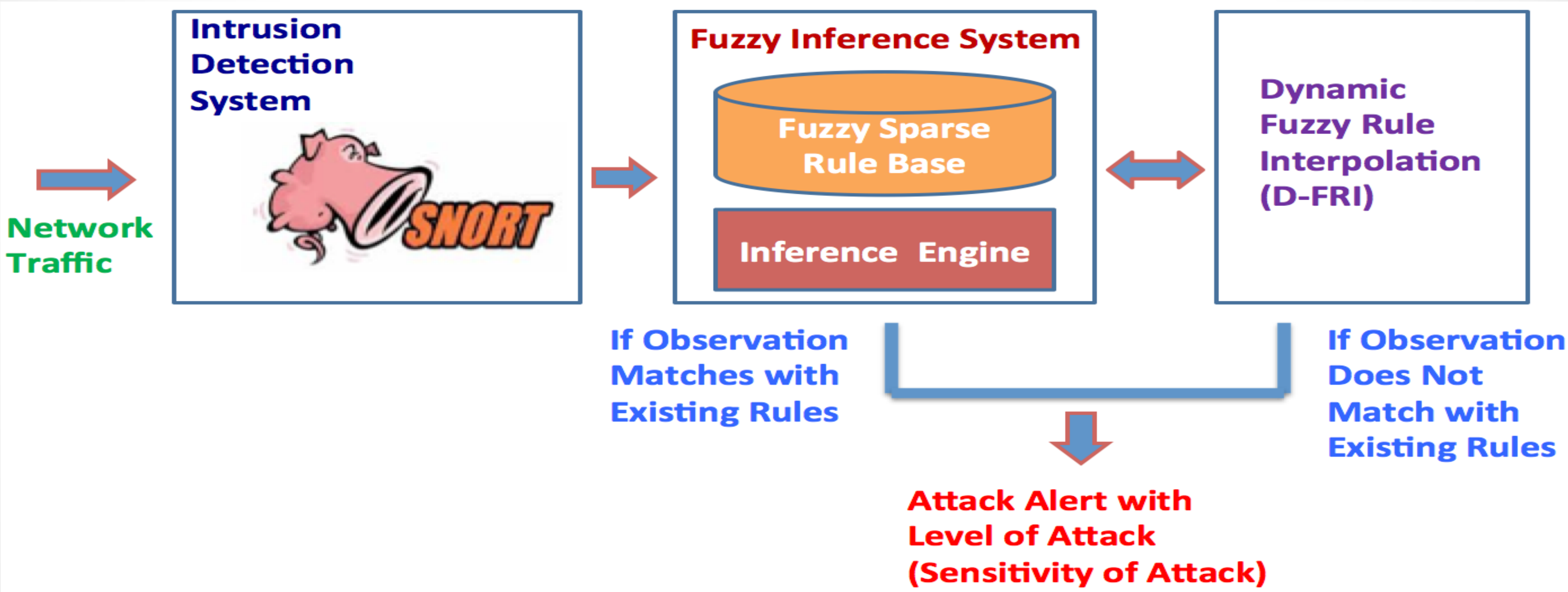
- Task: Protecting networked information system resources, by:
  - Preserving authorised restrictions on access to, and disclosure of, information
  - Guarding against unauthorised alternation to, or destruction of, information
  - Ensuring timely and reliable authorised access to, and use of, information

- Approach: Using intelligent systems to detect/prevent/recover from security attack (that compromises information security)

- Challenge: Knowledge being imprecise and sparse, and requirements changing over time

Network Intrusion Detection

- D-FRI-Snort: Enhances Snort with D-FRI
- Snort (open source IDS): collects and monitors network traffic data and generates attack alerts



Performance of D-FRI-Snort

- Snort vs D-FRI-Snort alert outputs

Obs.	Input			Output - Attack Alert	
No.	*ATP*	*NPS*	*NPR*	*Snort PSA*	*D-FRI-Snort PSA*
1	17.78	283	1167	no alert	very low attack alert
2	11.21	605	1764	no alert	low attack alert
3	8.03	1105	2506	no alert	medium attack alert
4	6.57	1317	3068	no alert	high attack alert
5	5.28	1642	3657	no alert	very high attack alert

ATP: Average time between received packets
NPS: Number of packets sent
NPR: Number of packets received
PSA: Port scan attack

Strengthening Results with Dynamic Learning

Accuracy with D-FRI-Snort dynamic rule promotion

	€%dvi	€%dvt	€%ivt
AVG	2.40	**1.31**	2.56
SD	2.72	**1.32**	2.68

D-FRI-Snort alert outputs after dynamic rule promotion

Obs.	Input			Attack Alert
No.	*ATP*	*NPS*	*NPR*	*D-FRI-Snort PSA*
1	6.95	1267	2385	high attack alert
2	5.23	643	1875	low attack alert
3	4.61	996	3010	high attack alert
4	7.91	1005	2805	medium attack alert
5	15.64	310	2266	low attack alert

Application: Mammogram Mass Analysis

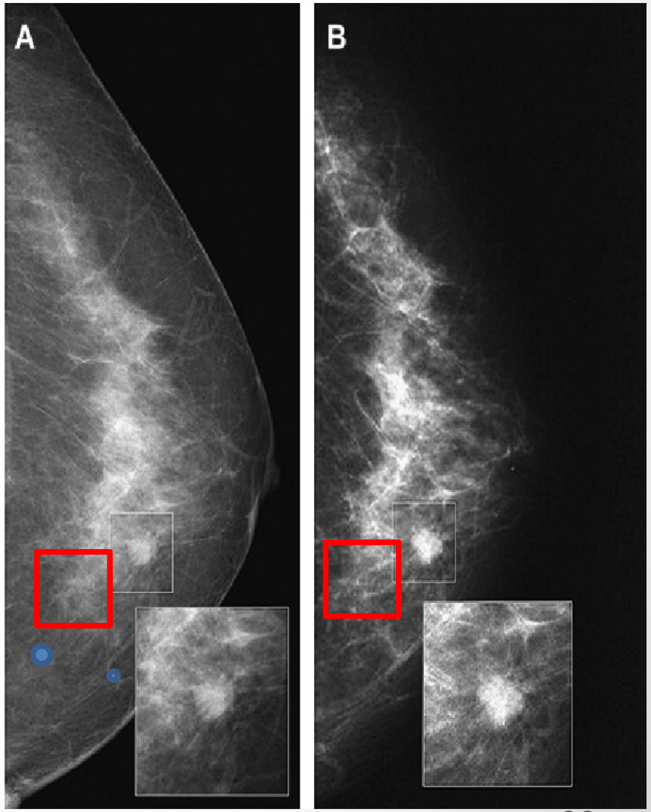
- Mammogram: Image obtained by mammography

- Mass: Group of cells clustered together more densely compared to surrounding tissues

- Benign/Malignant mass
  - Combination of *Shape*, *Margins* and *Density* --> Benign/Malignant

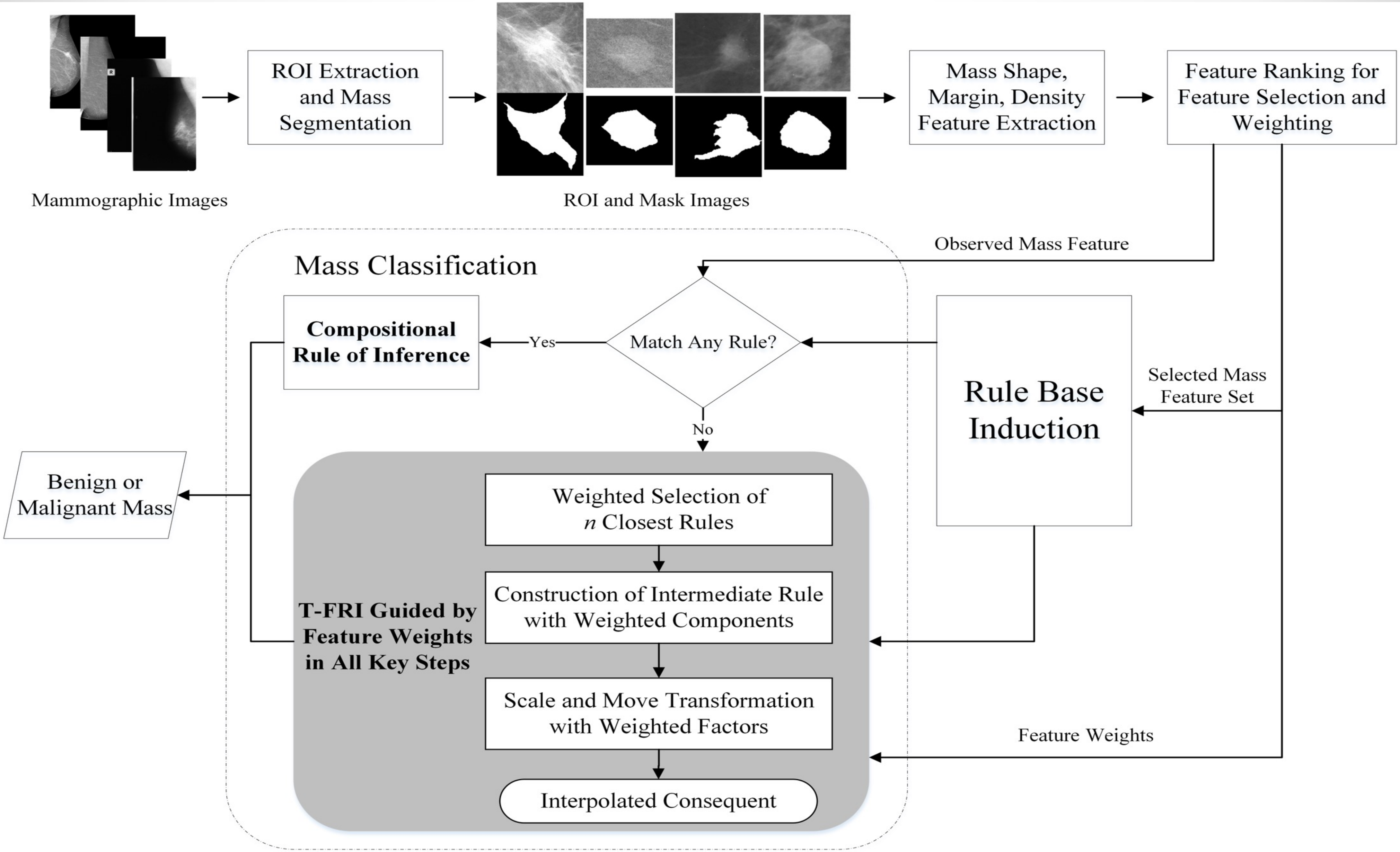
- Challenge: Insufficient coverage of problem space

Region of Interest (ROI)



Keyriläinen, Jani, et al. "Phase-contrast X-ray imaging of breast." Acta radiologica 51.8 (2010): 866-884.

Mammogram Mass Classification



Classification Performance

- Use of entire rule set learned from BCDR-D01/BCDR-F01 (10x10 CV)

BCDR-D01				
Schemes	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
CRI	83.44	78.57	86.59	-
T-FRI	91.22	87.85	93.41	0.9607
W-T-FRI	91.65	88.93	93.41	0.9614
BCDR-F01				
Schemes	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
CRI	83.73	81.30	86.27	-
T-FRI	84.28	82.14	86.49	0.9019
W-T-FRI	84.28	82.14	86.49	0.9023

- State-of-the-art (explainable): AUC = 0.9650 / 0.8940 for BCDR-D01 / BCDR-F01 *

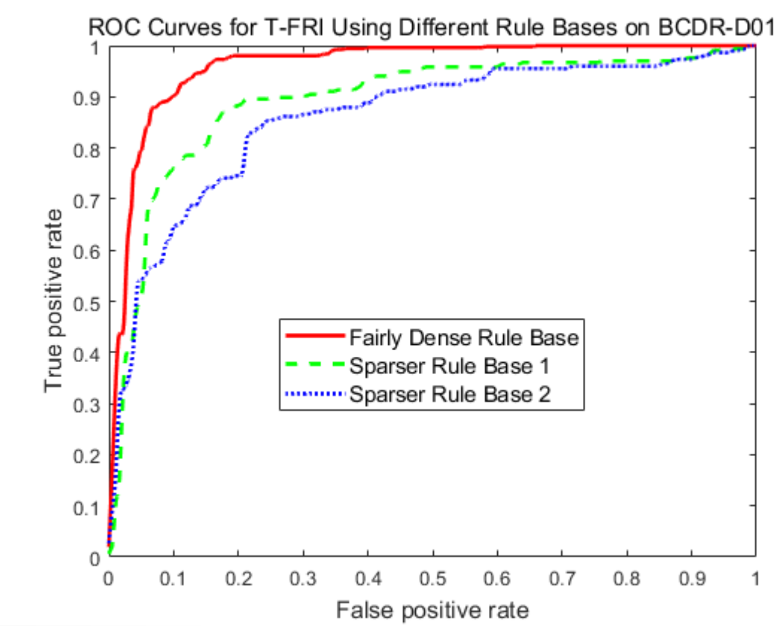
* D.C. Moura et.al "An evaluation of image descriptors combined with clinical data for breast cancer diagnosis." International journal of computer assisted radiology and surgery 8.4 (2013): 561-574.

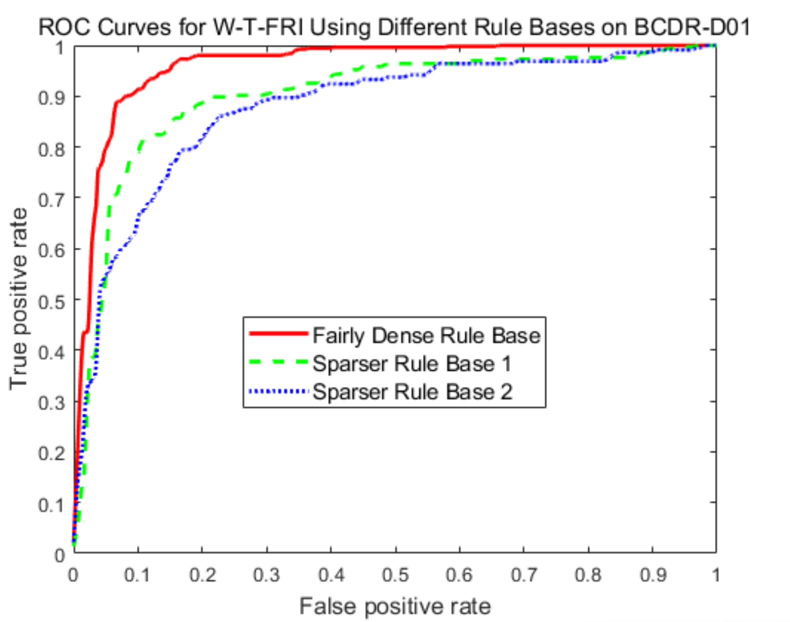
Classification Performance (Cont'd)

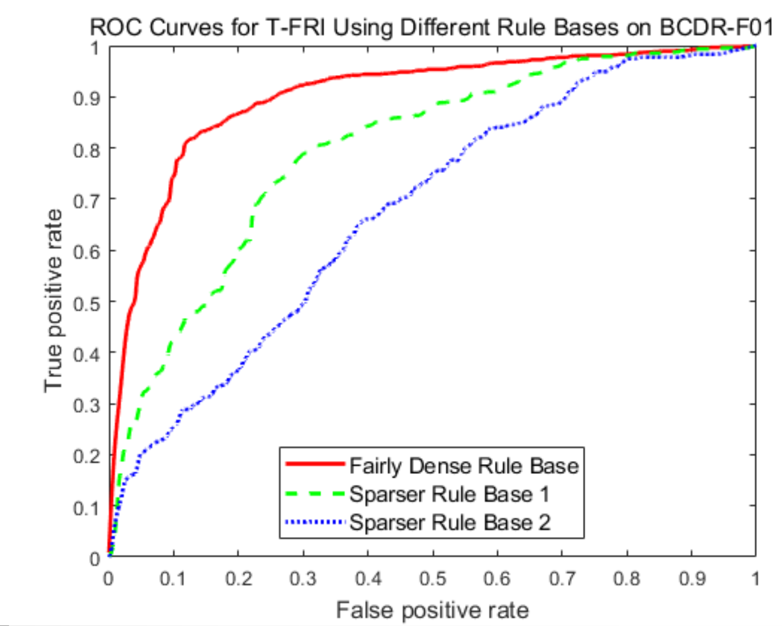
- Use of sparser rule bases (10x10 CV)

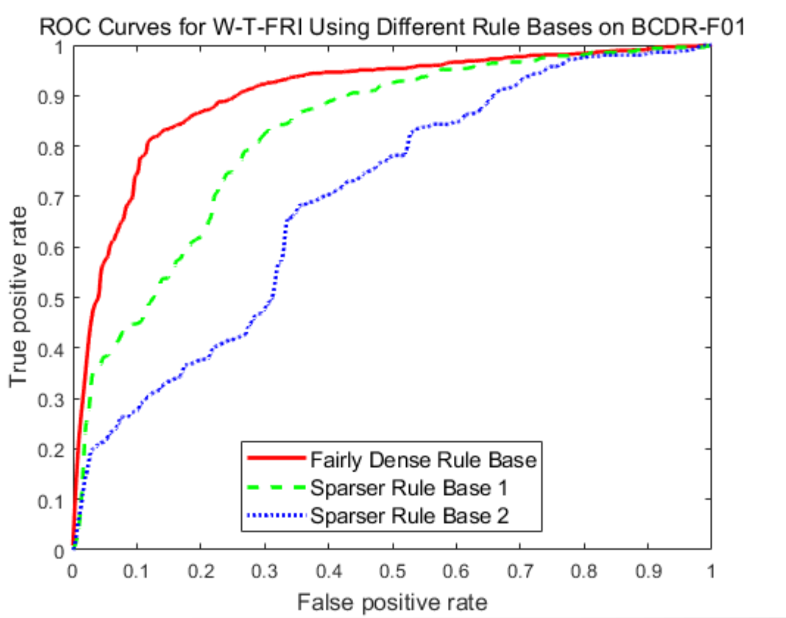
	Schemes	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
BCDR-D01	Sparser Rule Base 1 (30% removed)				
	CRI	53.43	44.04	59.60	-
	T-FRI	83.71	78.57	87.06	0.8918
	W-T-FRI	86.07	81.55	89.02	0.9010
	Sparser Rule Base 2 (70% removed)				
	Schemes	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
	CRI	23.44	12.95	30.29	-
	T-FRI	77.67	74.55	79.70	0.8589
	W-T-FRI	81.75	79.02	83.53	0.8784
BCDR-F01	Sparser Rule Base 1 (30% removed)				
	Schemes	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
	CRI	38.42	29.41	47.99	-
	T-FRI	73.47	71.22	75.89	0.7948
	W-T-FRI	75.09	73.95	76.34	0.8238
	Sparser Rule Base 2 (70% removed)				
	Schemes	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
	CRI	16.33	12.60	20.24	-
	T-FRI	62.60	63.30	61.91	0.6833
	W-T-FRI	66.33	68.06	64.58	0.7040

Classification Performance (Cont'd)









Conclusion

- AI can still work with limited (and imprecise) data
  - AI =\= Deep learning
  - T-FRI offers a potentially powerful approach for explainable AI
- Further extensions/improvements/alternatives
  - Closed form T-FRI of mathematical rigour
  - T-FRI with different distance functions
  - Weighted FRI for non transformation-based approaches
  - TSK/ANFIS model-based FRI
  - Integration of FRI and compositional rule of inference
  - Applications: counter-terrorism, image super-resolution
- On-going and future work
  - Vector form T-FRI of mathematical rigour
  - D-FRI with dynamic rule-pruning using FRI
  - T-FRI with different aggregation functions
  - FRI over a group of observations
  - Theoretical analysis of empirically revealed FRI properties
  - More real-world applications

Acknowledgement and References

- Thanks to former students who did the real work reported here:
  - Dr T. Chen, Prof C. Cheng, Dr R. Diao, Dr Z. Huang, Prof S. Jin, Dr F. Li, Dr N. Naik, Dr J. Yang, Prof L. Yang, Dr P. Zhang, Dr M. Zhou
- Selected references:

  - Transformation-based fuzzy rule interpolation with Mahalanobis distance measures supported by Choquet integral. ***IEEE Transactions on Fuzzy Systems***, 31(4):1083-1097, 2023.
  - W-Infer-polation: Approximate reasoning via integrating weighted fuzzy rule inference and interpolation. ***Knowledge-Based Systems***, 258:109995, 2022.
  - Fuzzy rule interpolation with k-neighbours for TSK models. ***IEEE Transactions on Fuzzy Systems***, 40(10):4031-4043, 2022.
  - Constructing ANFIS with sparse data through group-based rule interpolation: an evolutionary approach. ***IEEE Transactions on Fuzzy Systems***, 30(4):893-907, 2022.
  - Approximate reasoning with fuzzy rule interpolation: Background and recent advances. ***Artificial Intelligence Review***, 54:4543-4590, 2021.
  - ANFIS construction with sparse data via group rule interpolation. ***IEEE Transactions on Cybernetics***, 51(5):2773-2786, 2021.
  - D-FRI-Honeypot: a secure sting operation for hacking the hackers using dynamic fuzzy rule interpolation. ***IEEE Transactions on Emerging Topics in Computational Intelligence***, 5(6): 893-907, 2021.
  - Bidirectional approximate reasoning-based approach for decision support. ***Information Sciences***, 506:99-112, 2020.
  - Interpretable mammographic mass classification with fuzzy interpolative reasoning. ***Knowledge-Based Systems***, 191:105279, 2020.
  - Interpolation with just two nearest neighbouring weighted fuzzy rules. ***IEEE Transactions on Fuzzy Systems***, 28(9):2255-2262, 2020.
  - A new approach for transformation-based fuzzy rule interpolation. ***IEEE Transactions on Fuzzy Systems***, 28(12):3330-3344, 2020.
  - Fuzzy knowledge-based prediction through weighted rule interpolation. ***IEEE Transactions on Cybernetics***, 50(10):4508-4517, 2020.
  - Interpretable mammographic mass classification with fuzzy interpolative reasoning. ***Knowledge-Based Systems***, 191:105279, 2020.
  - Bidirectional approximate reasoning-based approach for decision support. ***Information Sciences***, 506:99-112, 2020.
  - Fuzzy rule-based interpolative reasoning supported by attribute ranking. ***IEEE Transactions on Fuzzy Systems***, 26(5): 2758-2773, 2018.
  - Dynamic fuzzy rule interpolation and its application to intrusion detection. ***IEEE Transactions on Fuzzy Systems***, *26(4):1878-1892, 2018*.
  - Generalised adaptive fuzzy rule interpolation. ***IEEE Transactions on Fuzzy Systems***, 2016.
  - Rough-fuzzy rule interpolation. ***Information Sciences***, 351:1-17, 2016.
  - Backward fuzzy rule interpolation. ***IEEE Transactions on Fuzzy Systems***, 22(6): 1682-1698, 2014.
  - Closed form fuzzy interpolation. ***Fuzzy Sets and Systems***, 225:1-22, 2013.
  - Adaptive fuzzy interpolation. ***IEEE Transactions on Fuzzy Systems***, 19(6):1107-1126, 2011.
  - Fuzzy interpolative and extrapolative reasoning: a practical approach. ***IEEE Transactions on Fuzzy Systems***, 16(1):13-28, 2008.
  - Fuzzy interpolative reasoning via scale and move transformation. ***IEEE Transactions on Fuzzy Systems***, 14(2):340-359, 2006.

Thank you!

Questions?