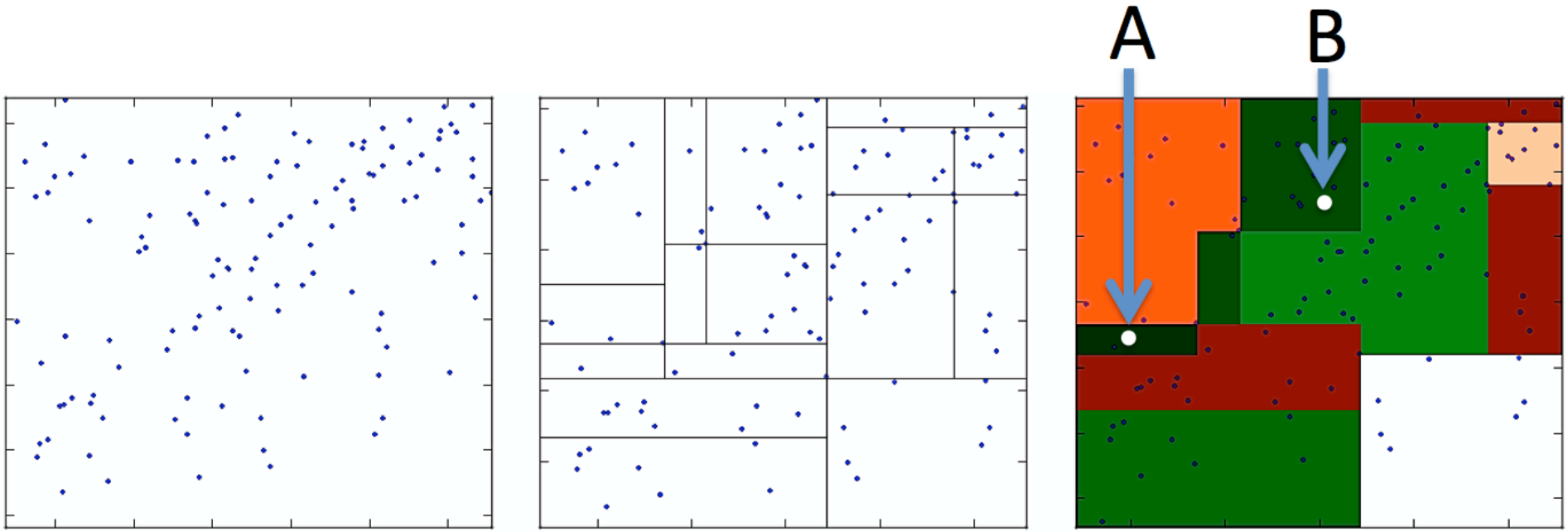
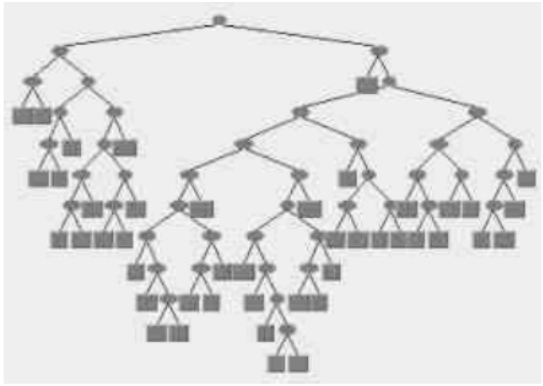
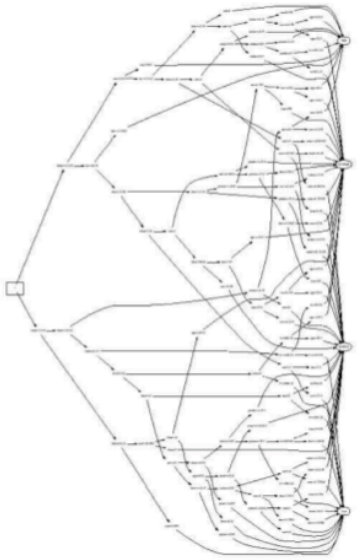
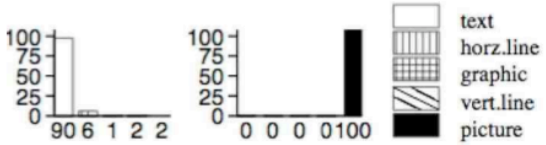
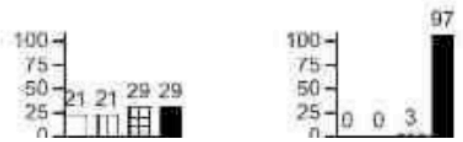
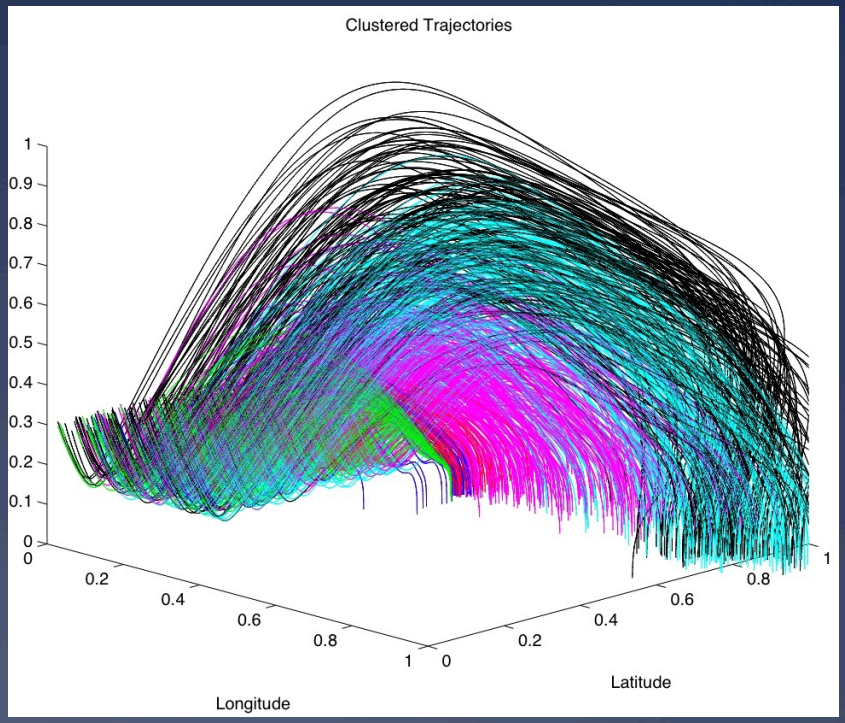

A
B

Contrast Set Learning ➔ Succinct Explanations

find graphics on a page from 11 features	find good housing in Boston
	
	
$34 \leq height < 86 \wedge$ $3.9 \leq mean\_tr < 9.5$	$6.7 \leq RM < 9.8 \wedge$ $12.6 \leq PTRATION < 15.9$

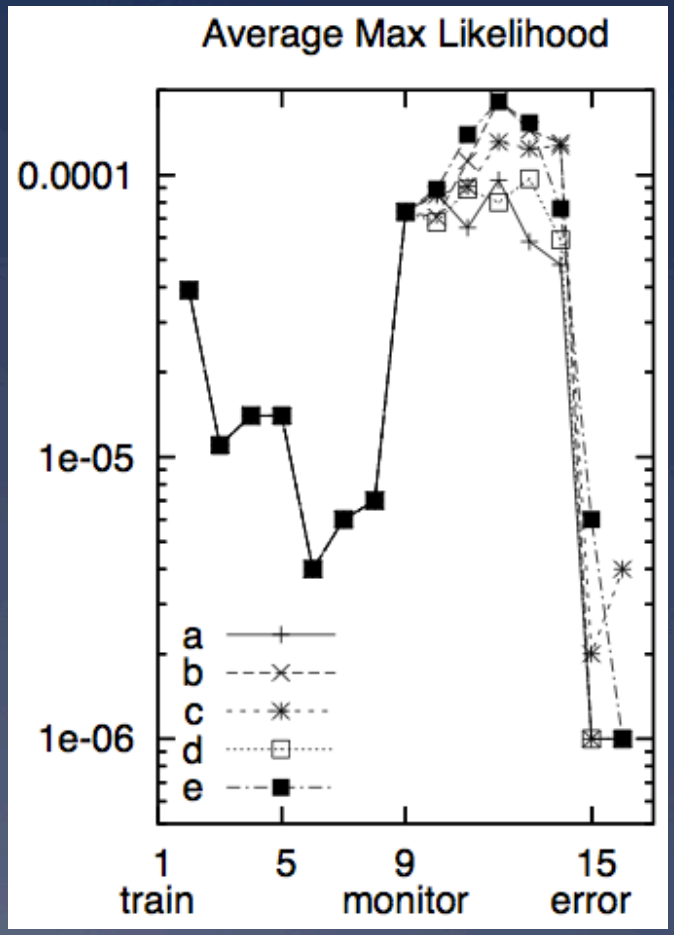
Contrast Set Learning (10 years later)

- No longer a post-processor to a decision tree learner
  - TAR3: Branch pruning operators applied directly to discretized data
- Summer'09
  - Shoot 'em up at NASA AMES
  - State-of-the-art numerical optimizer
  - TAR3
    - Ran 40 times faster
    - Generated better solutions
- Powerful succinct explanation tool



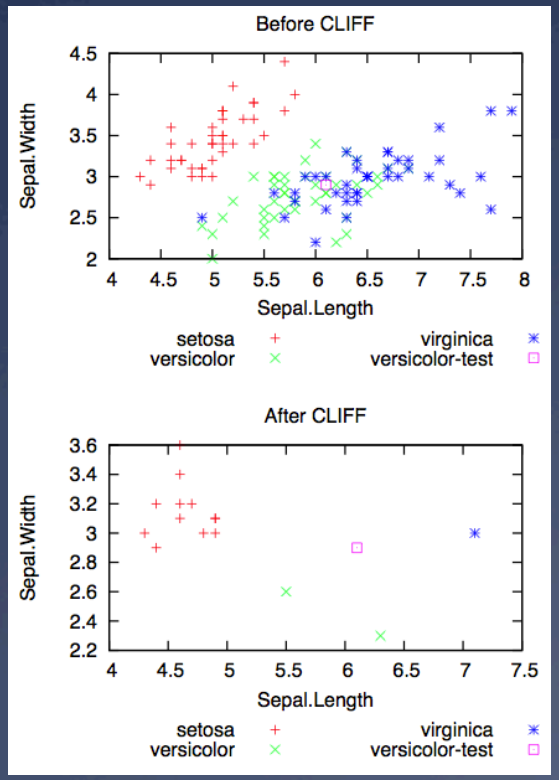
Contrast Set Learning ➔ Anomaly Detection

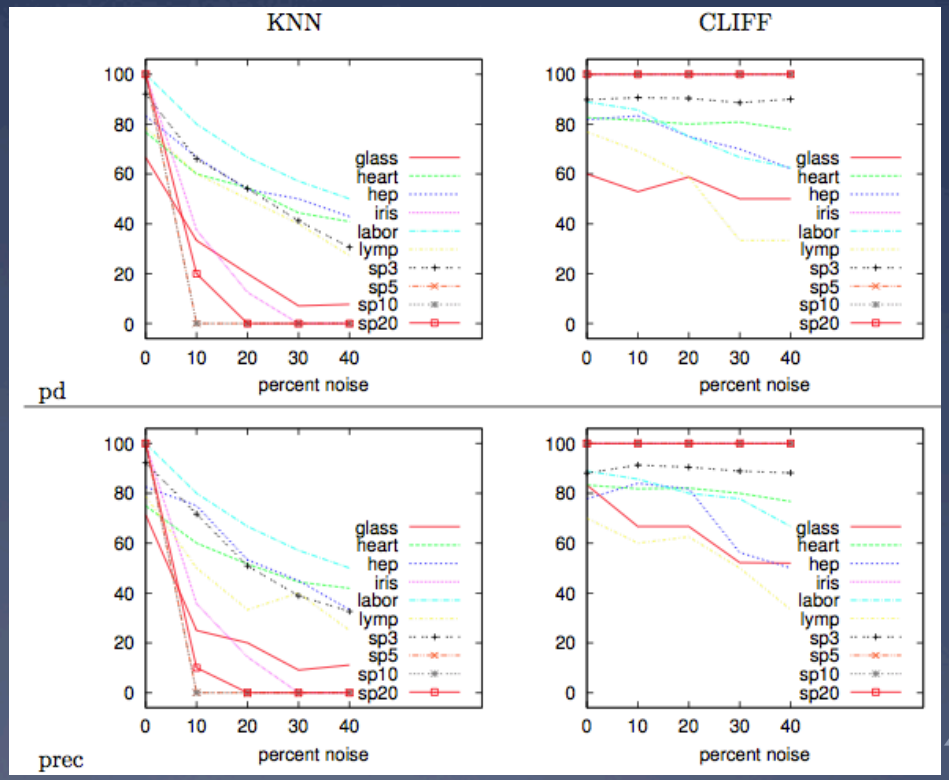
- Recognize when old ideas are now out-dated
- SAWTOOTH:
  - read data in "eras" of 100 instances
  - Classify all examples as "seen it"
- SAWTOOTH1:
  - Report average likelihood of examples belong to "seen it"
  - Alert if that likelihood drops
- SAWTOOTH2:
  - Back-end to TAR3
  - Track frequency of contrast sets
  - Some uniformity between contrast sets and anomaly detection



Contrast sets ➔ noise management

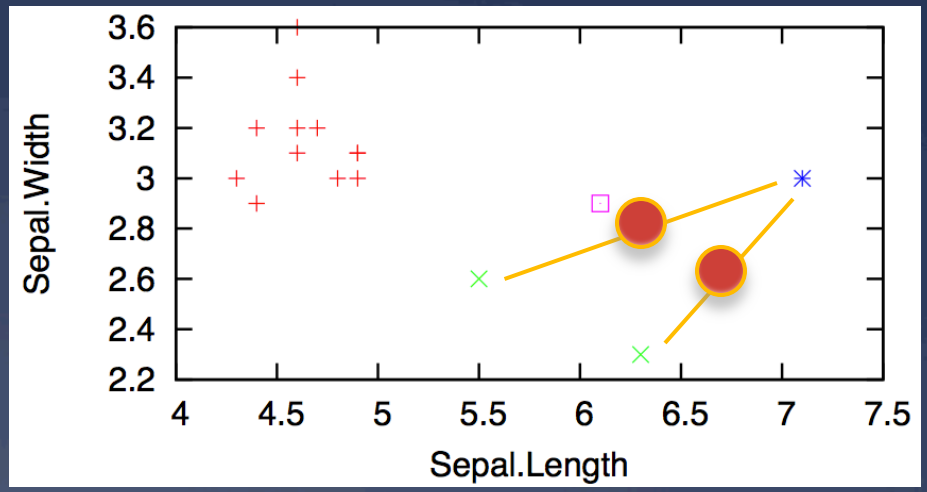
- CLIFF: post-processor to TAR3
  - Linear time instance selector
- Finds the attribute ranges that change classification
- Delete all instances that lack the "power ranges"





Contrast Sets ➔ CLIFF ➔ Active Learning

- Many examples, too few labels
- Reduce the time required for business users to offer judgment on business cases
- Explore the reduced space generated by CLIFF.
  - Randomly ample the instances half-way between different classes
- Fast (in the reduced space)



Contrast sets ➔ CLIFF ➔ Statistical databases

- Anonymize the data: Preserving its distributions
- For KNN, that means keep the boundaries between classes
  - Which we get from CLIFF
- Also, CLIFF empties out the instance space
  - Leaving us space to synthesize new instances