



TITCH:
**Attribute selection based on
discrimination power and frequency**

Philipp Spanger, Kurosawa Takahiro, Tokunaga Takenobu

*Department of Computer Science
Tokyo Institute of Technology*

Two factors in attribute selection

- Human's general preference on object attributes
e.g. colour > orientation
→ Frequency of attributes
- Salience of attributes in a specific case
(particular situation in a domain)
→ Discrimination power

Base algorithm

(Using case specific attribute salience)

- Ranking attribute-value pairs according to their discrimination power
 - Disc. power = No. of objects excluded by specifying an attr.-value pair
- Selecting attributes one-by-one until unique identification of the target

Domain	Dice	
	static	dynamic
Furniture	0.588	0.601
People	0.559	0.559

Improvement of base algorithm

(Introducing Human's general preference on attr.)

- Human's general preference on attributes \cong Frequency of attributes in the domain
- Assumption: Salient attributes tend to be used frequently
- Weighting attributes based on frequency of occurrence in the data (absolute, relative)

Domain	Dice		
	Base	Abs.	Rel.
Furniture	0.588	0.685	0.707
People	0.559	0.651	0.648
People+	---	0.683	0.678

Indispensable features

- “Type”
 - Used all instances in both domains
- Dependency between attributes
 - hairColour → hasHair, hasBeard
- What’s common in above two cases?
 - Those could be realised as a head of NP

Difficulty of cases

- Metric representing a difficulty of a given case?
- No. of possible different sets of attr.-value pairs uniquely identifying the target
- But, no correlation the success of our algorithm
- This reflects the fact that humans select attributes from a very limited set of combinations, independent of the search space