UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

PREFERENCE LEARNING: MACHINE LEARNING MEETS PREFERENCE MODELING

Eyke Hüllermeier
Intelligent Systems Group
Department of Computer Science
University of Paderborn, Germany

eyke@upb.de

PREFERENCES ARE UBIQUITOUS

INTELLIGENT SYSTEMS

Preferences play a key role in many applications of computer science and modern information technology:

COMPUTATIONAL ADVERTISING	RECOMMENDER SYSTEMS	COMPUTER GAMES
AUTONOMOUS AGENTS	ELECTRONIC COMMERCE	ADAPTIVE USER INTERFACES
PERSONALIZED MEDICINE	ADAPTIVE RETRIEVAL SYSTEMS	SERVICE-ORIENTED COMPUTING

PREFERENCES ARE UBIQUITOUS

INTELLIGENT SYSTEMS

Preferences play a key role in many applications of computer science and modern information technology:

COMPUTATIONAL ADVERTISING	RECOMMENDER SYSTEMS	COMPUTER GAMES
AUTONOMOUS AGENTS	ELECTRONIC COMMERCE	ADAPTIVE USER INTERFACES
PERSONALIZED MEDICINE	ADAPTIVE RETRIEVAL SYSTEMS	SERVICE-ORIENTED COMPUTING

medications or therapies specifically tailored for individual patients

COMMERCIAL INTEREST

Amazon files patent for “anticipatory” shipping

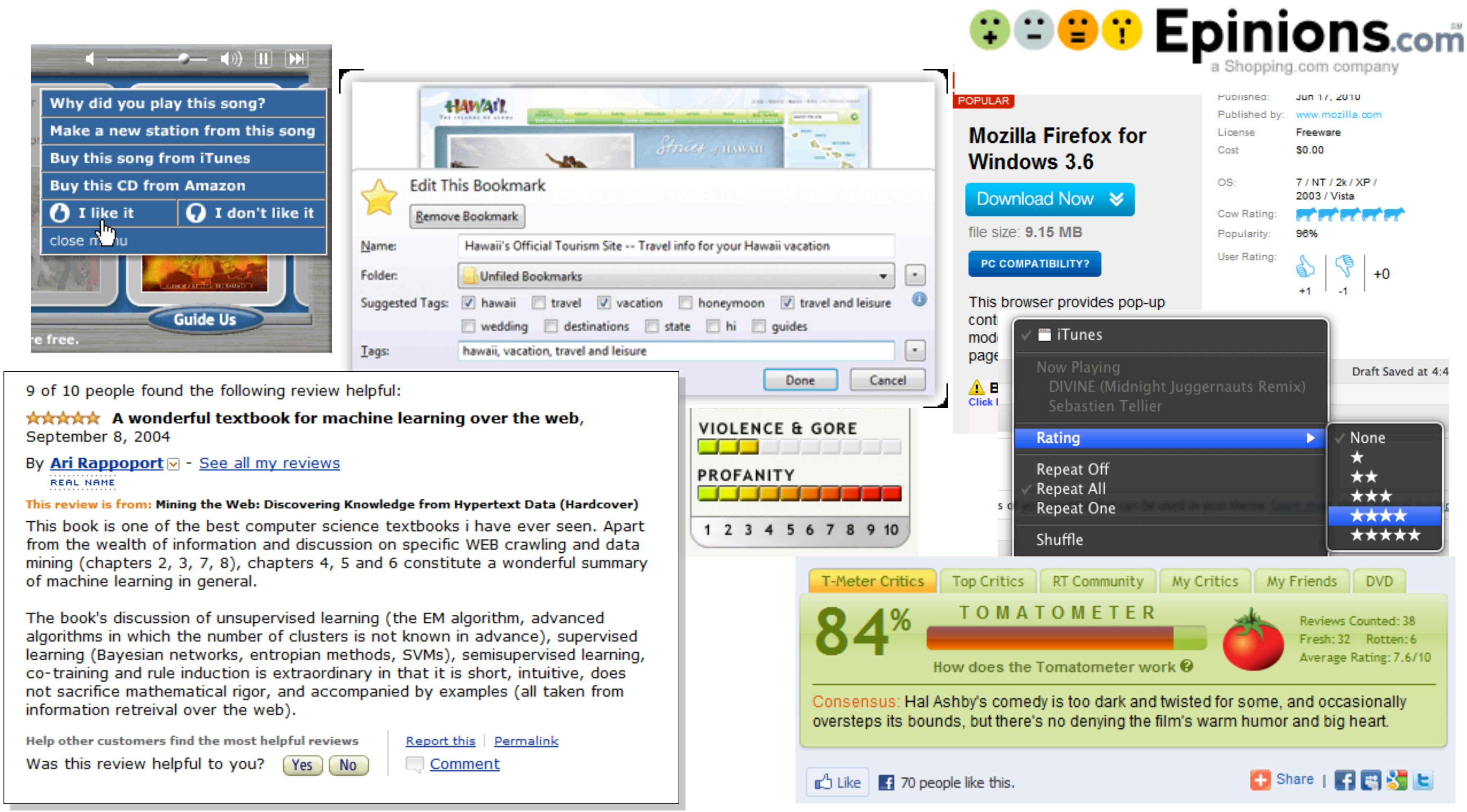
10 Comments / Shares / Tweets / Stumble / Email / More +

Amazon.com has filed for a patent for a shipping system that would anticipate what customers buy to decrease shipping time.

Amazon says the shipping system works by analyzing customer data like, purchasing history, product searches, wish lists and shopping cart contents, the Wall Street Journal reports. According to the patent filing, items would be moved from Amazon’s fulfillment center to a shipping hub close to the customer in anticipation of an eventual purchase.

PREFERENCE INFORMATION

INTELLIGENT SYSTEMS



PREFERENCE INFORMATION

INTELLIGENT SYSTEMS

| Offizielle Homepage | **Daniel Baier** |
www.**daniel-baier**.com/
Willkommen auf der offiziellen Homepage von Fussballprofi **Daniel Baier** - TSV 1860 München.

NOT CLICKED ON

Prof. Dr. **Daniel Baier** - Brandenburgische Technische Universität ...
www.tu-cottbus.de/fakultaet3/de/.../team/.../prof-dr-**daniel-baier**.html
Vökler, Sascha; Krausche, **Daniel**; **Baier**, Daniel: Product Design Optimization Using Ant Colony And Bee Algorithms: A Comparison, erscheint in: Studies in ...

CLICKED ON

Daniel Baier
www.weltfussball.de/spieler_profil/**daniel-baier**/
Daniel Baier - FC Augsburg, VfL Wolfsburg, VfL Wolfsburg II, TSV 1860 München.

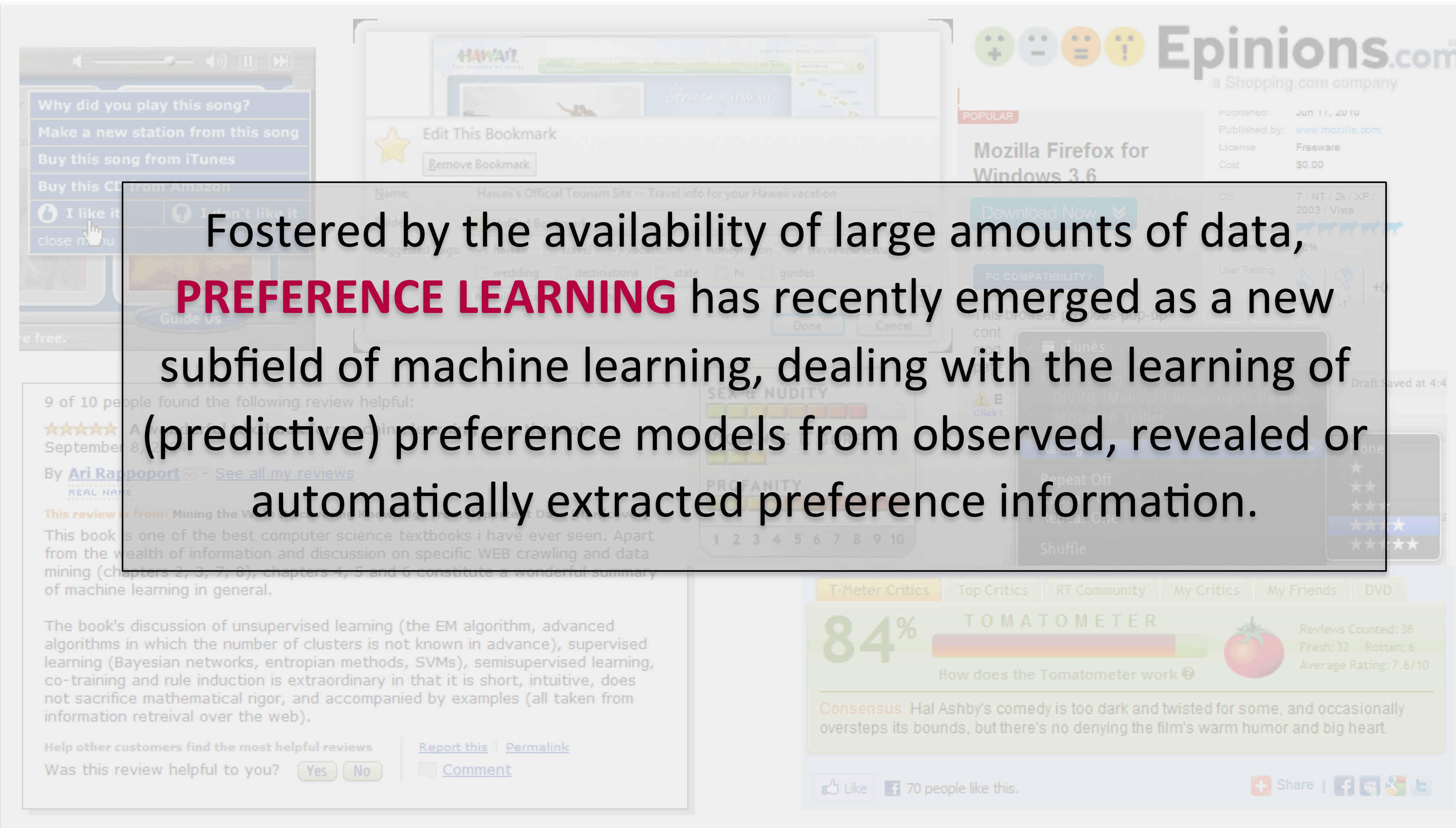
Daniel Baier - aktuelle Themen & Nachrichten - sueddeutsche.de
www.sueddeutsche.de/thema/**Daniel_Baier**
Aktuelle Nachrichten, Informationen und Bilder zum Thema **Daniel Baier** auf sueddeutsche.de.

Daniel Baier | Facebook
de-de.facebook.com/**daniel.baier**.589
Tritt Facebook bei, um dich mit **Daniel Baier** und anderen Nutzern, die du kennst, zu vernetzen. Facebook ermöglicht den Menschen das Teilen von Inhalten mit ...

FC Augsburg: Mein Tag in Bad Gögging: **Daniel Baier**
www.fcaugsburg.de/cms/website.php?id=/index/aktuell/news/...
2. Aug. 2012 – **Daniel Baier** berichtet heute, was für die Profis auf dem Programm stand. Hi FCA- Fans,. heute liegen wieder zwei intensive Trainingseinheiten ...

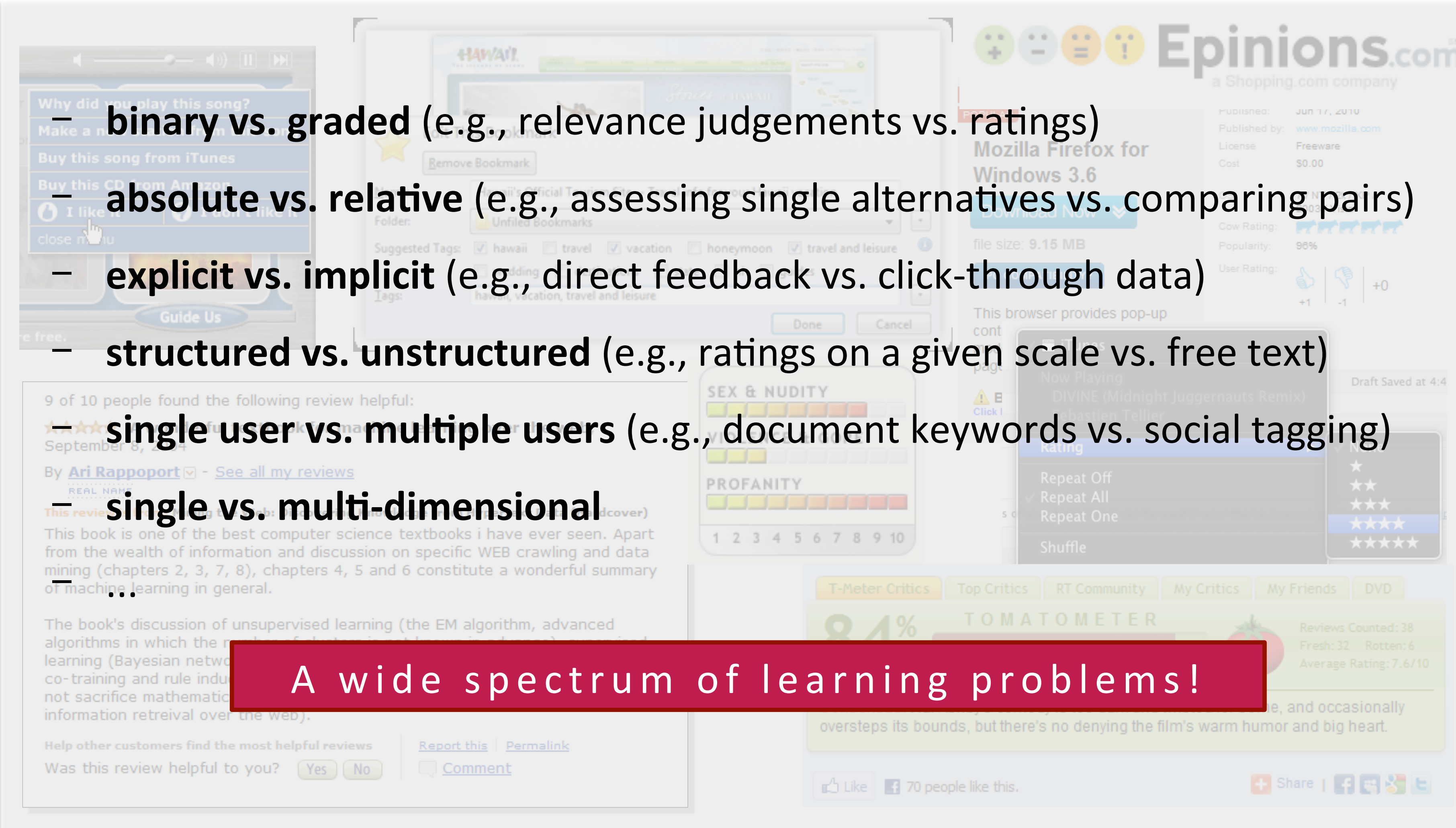
- *Preferences are not necessarily expressed explicitly, but can be extracted **implictly from people's behavior**!*
- *Massive amounts of very **noisy data**!*

PREFERENCE LEARNING



Fostered by the availability of large amounts of data, **PREFERENCE LEARNING** has recently emerged as a new subfield of machine learning, dealing with the learning of (predictive) preference models from observed, revealed or automatically extracted preference information.

TYPES OF PREFERENCES



- binary vs. graded (e.g., relevance judgements vs. ratings)
- absolute vs. relative (e.g., assessing single alternatives vs. comparing pairs)
- explicit vs. implicit (e.g., direct feedback vs. click-through data)
- structured vs. unstructured (e.g., ratings on a given scale vs. free text)
- single user vs. multiple users (e.g., document keywords vs. social tagging)
- single vs. multi-dimensional
- ...

A wide spectrum of learning problems!

PREFERENCE LEARNING TASKS

Preference learning problems are challenging, because

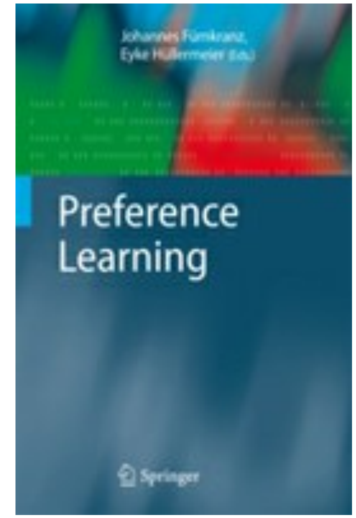
- sought predictions are complex/structured, *top-K ranking*
- supervision is weak (partial, noisy, ...), *clickthrough data*
- performance metrics are hard to optimize, *NDCG@K*
- ...

PL IS AN ACTIVE FIELD

INTELLIGENT SYSTEMS

Tutorials:

- European Conf. on Machine Learning, 2010
- Int. Conf. Discovery Science, 2011
- Int. Conf. Algorithmic Decision Theory, 2011
- European Conf. on Artificial Intelligence, 2012
- Int. Conf. Algorithmic Learning Theory, 2014



J. Fürnkranz &
E. Hüllermeier (eds.)
Preference Learning
Springer-Verlag 2011

Special Issue on Representing, Processing, and Learning Preferences: Theoretical and Practical Challenges (2011)

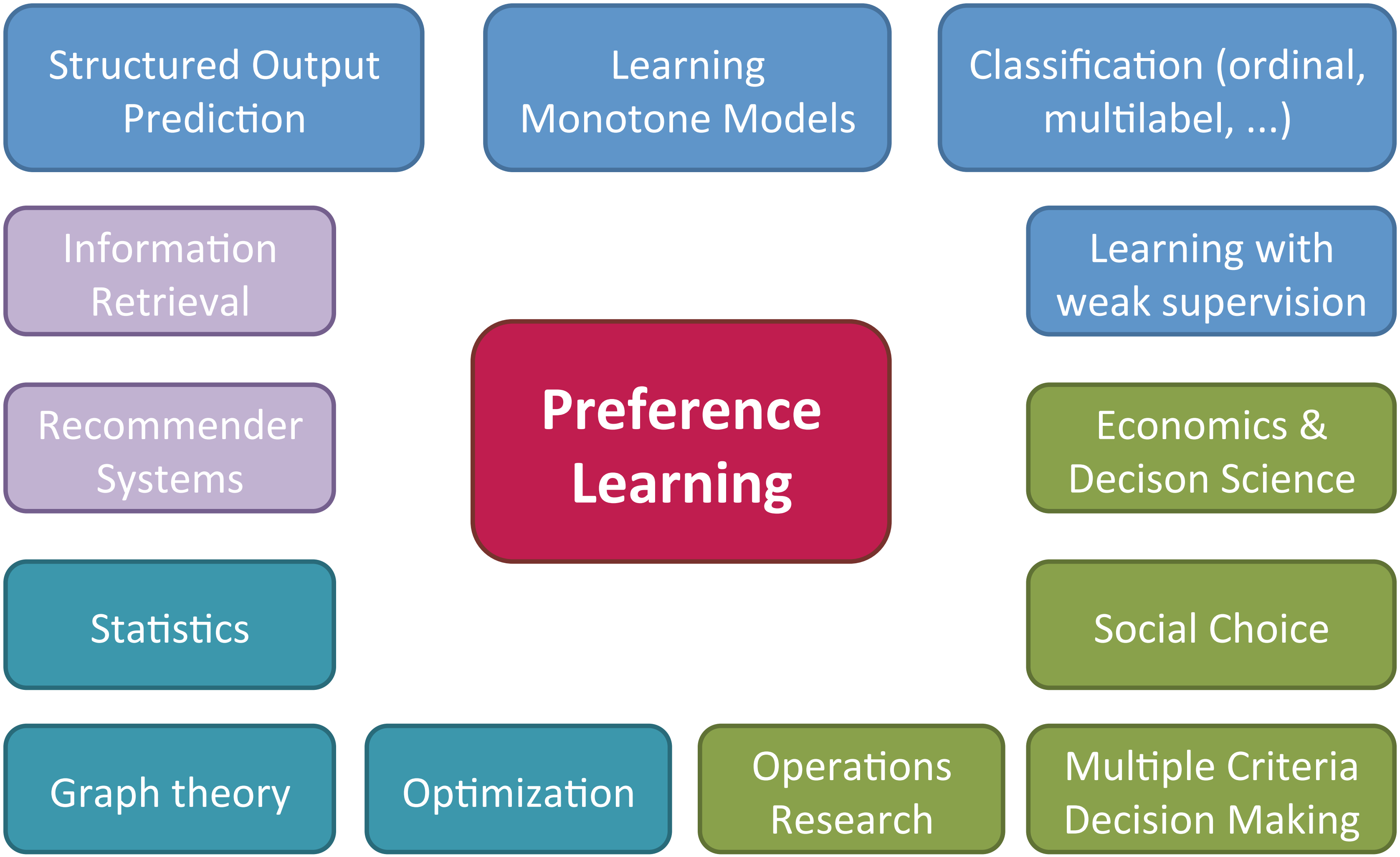
Special Issue on Preference Learning
Forthcoming

PL IS AN ACTIVE FIELD

INTELLIGENT SYSTEMS

- NIPS 2001: New Methods for Preference Elicitation
- NIPS 2002: Beyond Classification and Regression: Learning Rankings, Preferences, Equality Predicates, and Other Structures
- KI 2003: Preference Learning: Models, Methods, Applications
- NIPS 2004: Learning with Structured Outputs
- NIPS 2005: Workshop on Learning to Rank
- IJCAI 2005: Advances in Preference Handling
- SIGIR 07–10: Workshop on Learning to Rank for Information Retrieval
- ECML/PDKK 08–10: Workshop on Preference Learning
- NIPS 2009: Workshop on Advances in Ranking
- American Institute of Mathematics Workshop in Summer 2010: The Mathematics of Ranking
- NIPS 2011: Workshop on Choice Models and Preference Learning
- EURO 2009-12: Special Track on Preference Learning
- ECAI 2012: Workshop on Preference Learning: Problems and Applications in AI
- DA2PL 2012: From Decision Analysis to Preference Learning
- Dagstuhl Seminar on Preference Learning (2014)
- NIPS 2014: Analysis of Rank Data: Confluence of Social Choice, Operations Research, and Machine Learning

CONNECTIONS TO OTHER FIELDS

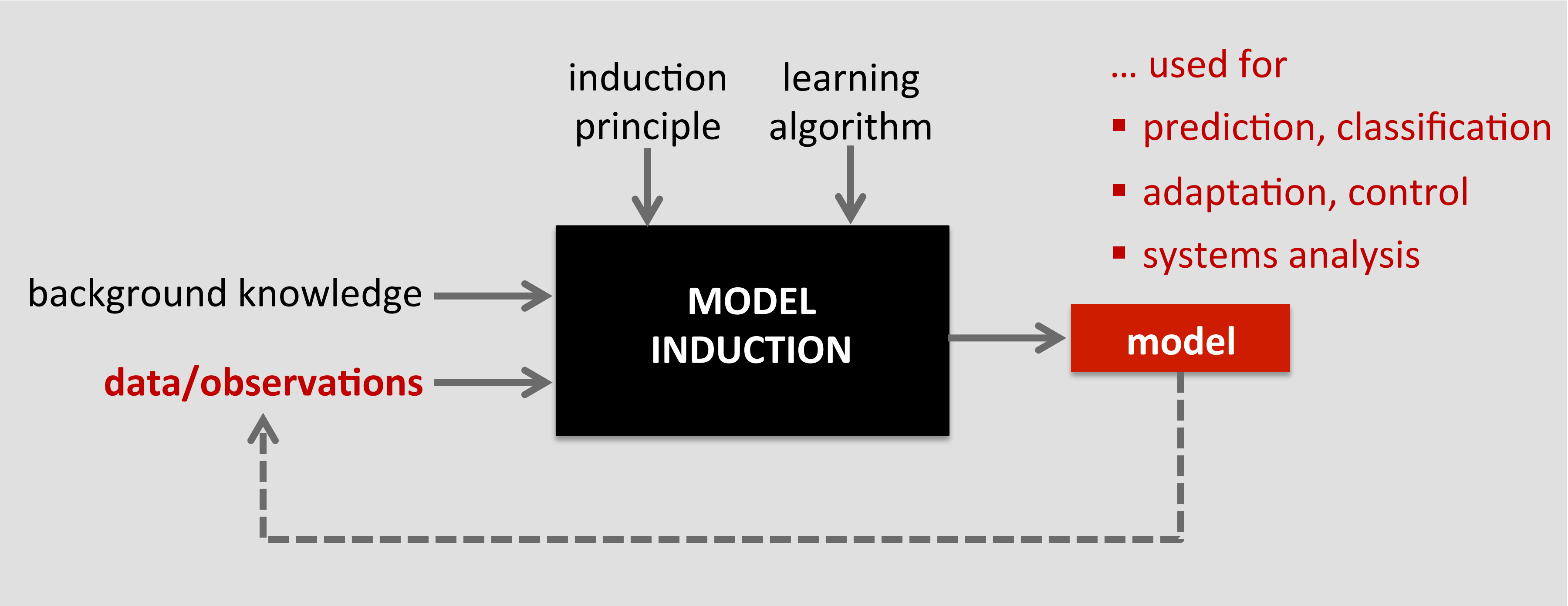


OUTLINE

INTELLIGENT SYSTEMS

PART 1	PART 2
Introduction to preference learning	Machine learning vs. MCDA
PART 3	**PART 4**
Multi-criteria preference learning	Preference-based online learning

SUPERVISED LEARNING: Algorithms and methods for discovering dependencies and regularities in a domain of interest, expressed through appropriate models, from specific observations or examples.



SUPERVISED LEARNING

INTELLIGENT SYSTEMS

	X_1	X_2	X_3	X_4	obs
TRAINING	0.34	0	10	174	***
	1.45	0	32	277	*
	1.22	1	46	421	****
	0.74	1	25	165	***
	0.95	1	72	273	*****
	1.04	0	33	158	***
TEST	0.85	0	45	194	?
	0.57	1	65	403	?
	1.32	1	26	634	?
	...	...	...	...	?

$$M : X_1 \times X_2 \times X_3 \times X_4 \longrightarrow \{*, \cdots, *****\}$$

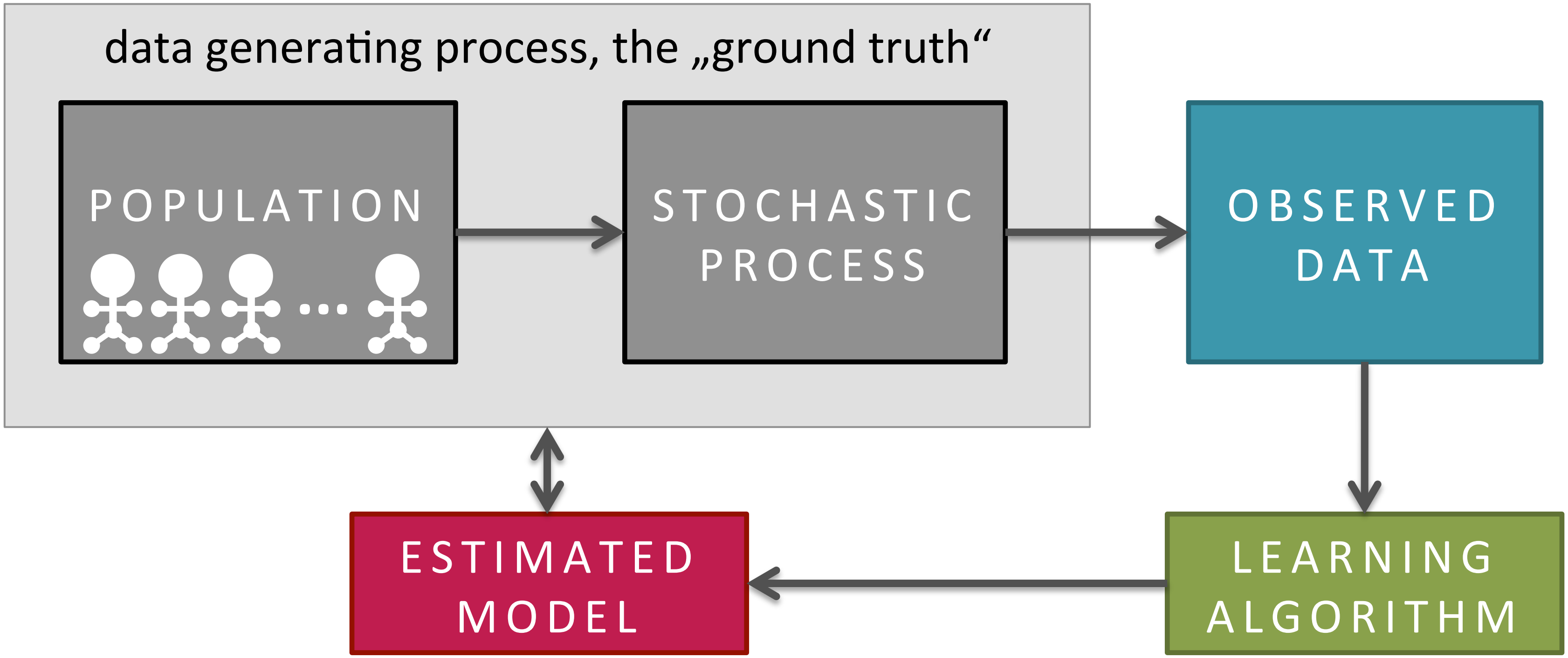
SUPERVISED LEARNING

INTELLIGENT SYSTEMS

	X_1	X_2	X_3	X_4	obs	pred
TRAINING	0.34	0	10	174	***	**
	1.45	0	32	277	*	*
	1.22	1	46	421	****	****
	0.74	1	25	165	***	****
	0.95	1	72	273	*****	****
	1.04	0	33	158	***	**
TEST	0.85	0	45	194	?	***
	0.57	1	65	403	?	*
	1.32	1	26	634	?	*****
	...	...	...	...	?	...

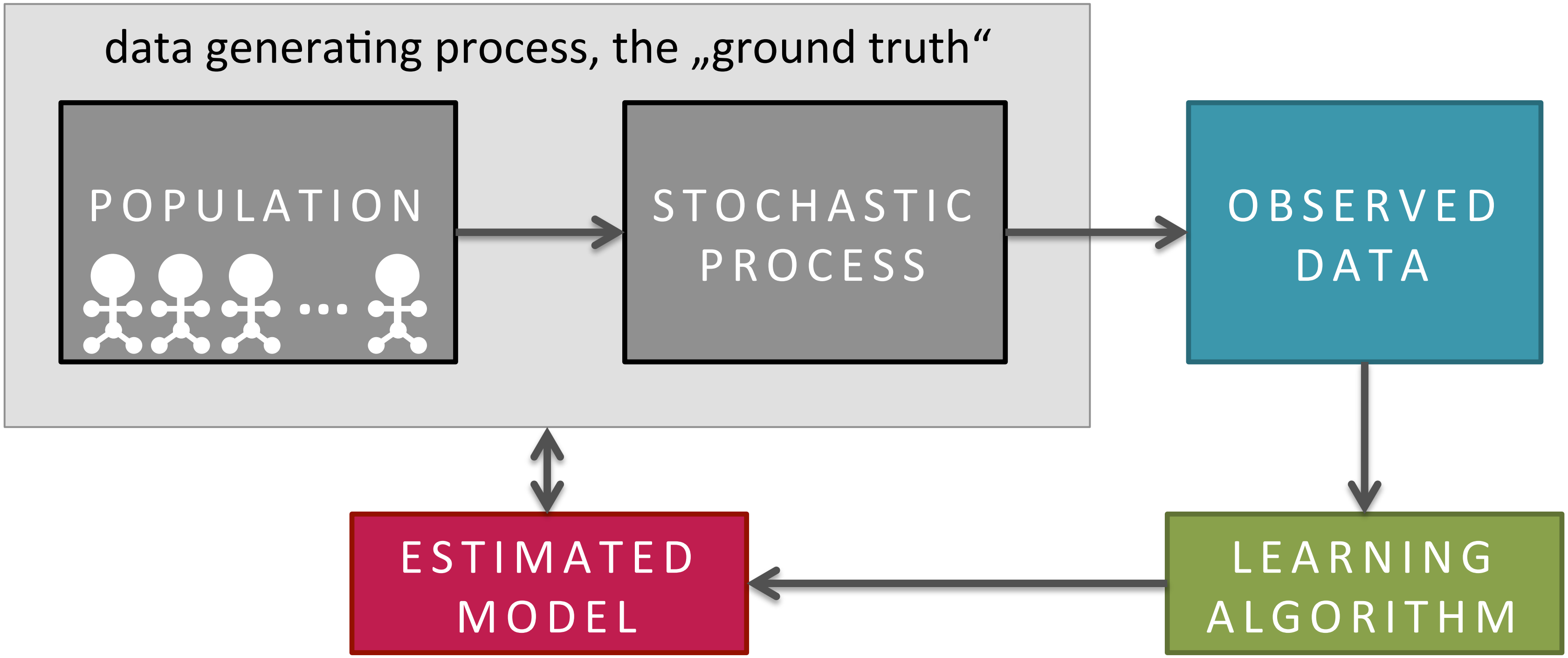
$$M : X_1 \times X_2 \times X_3 \times X_4 \longrightarrow \{*, \cdots, *****\}$$

MODEL INDUCTION IN ML



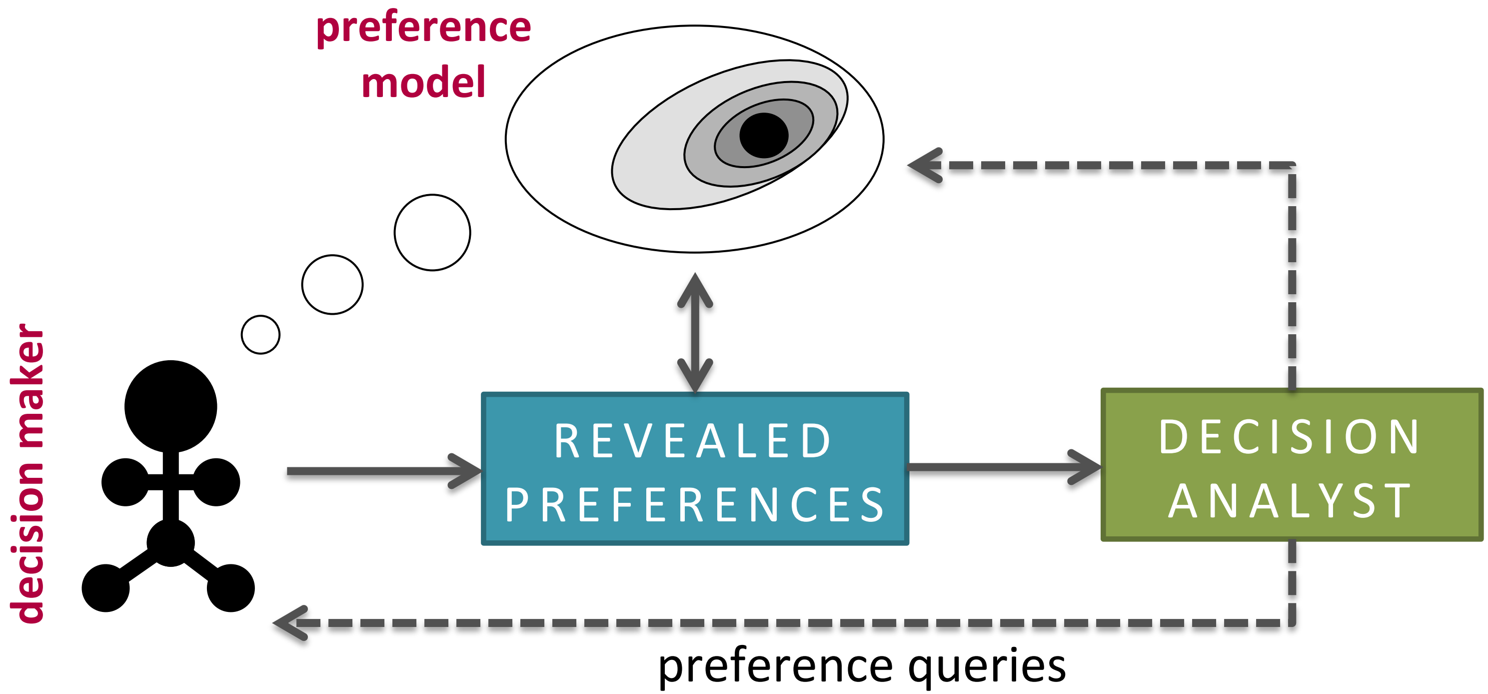
- *The model refers to an underlying **population of individuals**.*
- *Knowing the model allows for making good predictions **on average**.*
- *The **dependency** tried to be captured by the model is **not deterministic** (variability due to aggregation, „noisy" observations, etc.)*

MODEL INDUCTION IN ML



- *Assumptions about the „ground truth" allow for deriving **theoretical results** (given enough data, the learner is likely to get close to the target).*
- *Focus on predictive accuracy allows for simple **empirical comparison**.*

PREFERENCE ELICITATION IN MCDA



- *Single user, interactive process.*
- *Inconsistencies can be discovered and corrected.*
- *Constructive process, no „ground truth", no true vs. estimated model (construction vs. induction).*

MACHINE LEARNING VS. MCDA

INTELLIGENT SYSTEMS

This comparison is obviously simplified ....

- Other settings in ML: semi- and unsupervised learning, active learning, online learning, reinforcement learning, ...
- Bayesian approaches to preference elicitation (e.g. Viappini and Boutilier): stochastic setting, active learning using expected value of information.
- Machine learning with humans in the loop (Joachims): focusing on the interface between the human and a continuously learning system (beyond mere labeling).
- ...

OUTLINE

INTELLIGENT SYSTEMS

PART 1

Introduction to preference learning

PART 2

Machine learning vs. MCDA

PART 3

Multi-criteria preference learning

PART 4

Preference-based online learning

MULTI-CRITERIA PREFERENCE LEARNING

INTELLIGENT SYSTEMS

PREFERENCE MODELING

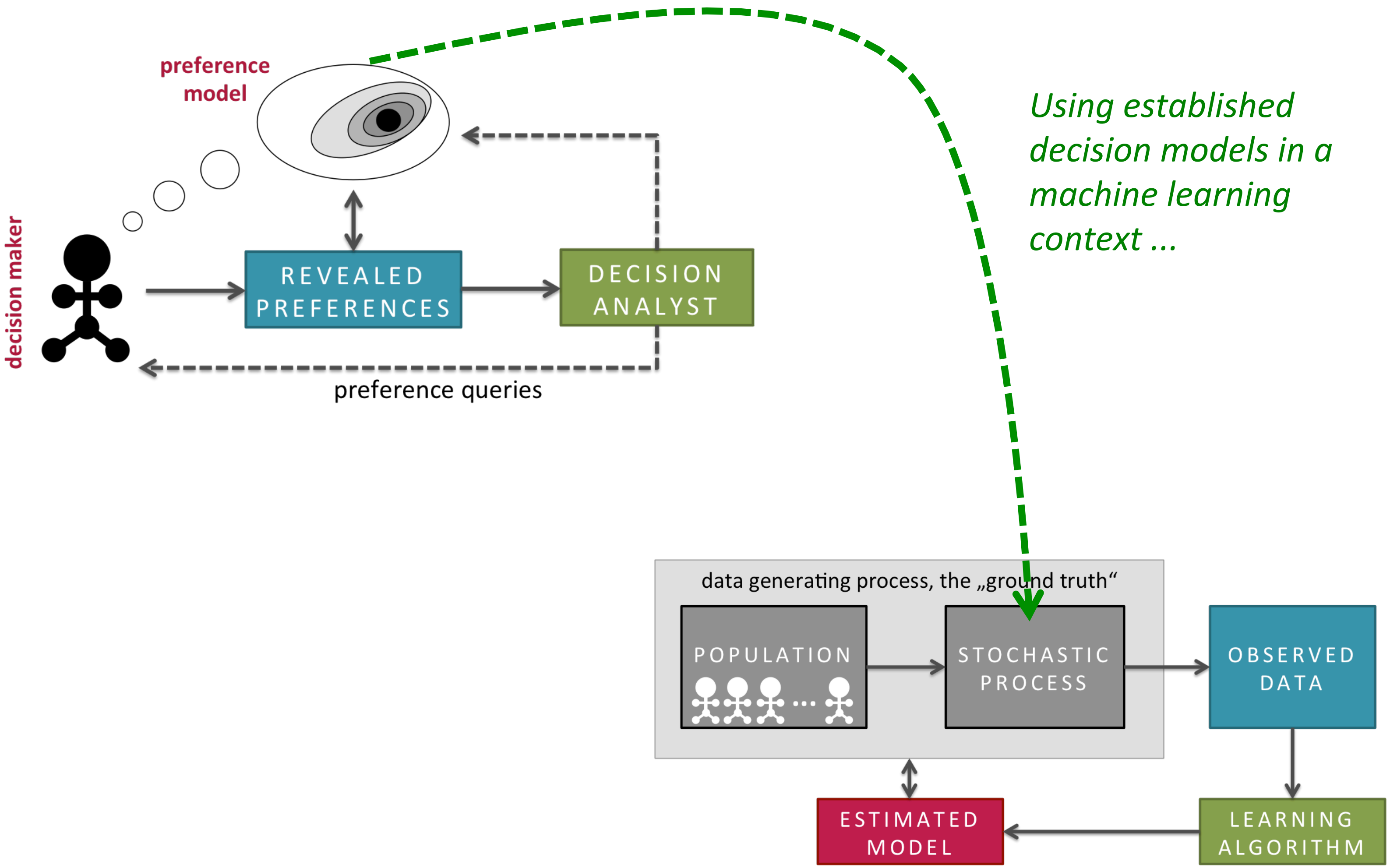
How do value and ranking functions generalize, what is their pedictive accuracy?

- **Choquet integral**
  [Fallah Tehrani et al., 2011, 2012, 2013]
- **CP nets, GAI nets**
  [Chevaleyre et al., 2010]
- **Lexicographic orders**
  [Bräuning and EH, 2012]
- **Majority rule models, MR Sort, ...**
  [Leroy et al., 2011; Sobrie et al., 2013]
- **TOPSIS-like models**
  [Agarwal et al., 2014]

How to represent value and ranking functions, and what properties should they obey?

PREFERENCE LEARNING

MULTI-CRITERIA PREFERENCE LEARNING



Using established decision models in a machine learning context ...

AGGREGATION OF CRITERIA

criteria
attributes/features

	Math	CS	Statistics	English	score
$\boldsymbol{x}_1$	16	17	12	19	0.7
$\boldsymbol{x}_2$	10	12	18	9	0.4
$\boldsymbol{x}_3$	19	10	18	10	0.6
	...	...	...	...	...
$\boldsymbol{x}_n$	8	18	10	18	0.5

Going beyond simple averaging ...

INTELLIGENT SYSTEMS

Non-additive measure (capacity) $\mu : 2^X \rightarrow [0,1]$:

We require ...

- $\mu(\emptyset) = 0$, $\mu(X) = 1$ — *normalization*
- $\mu(A) \leq \mu(B)$ for $A \subset B \subseteq X$ — *monotonicity*
- ~~$\mu(A \cup B) = \mu(A) + \mu(B)$ for $A \cap B = \emptyset$~~ — *not necessarily additivity*

NON-ADDITIVE MEASURES

Non-additive measures allow for modeling **interaction** between criteria:

- Positive (synergy): $\mu(A \cup B) > \mu(A) + \mu(B)$
- Negative (redundancy): $\mu(A \cup B) < \mu(A) + \mu(B)$

In a machine learning context: criteria = attributes/features

$\mu(A) =$ **joint importance** of the feature subset A

$\neq$ sum of individual importance degrees

DISCRETE CHOQUET INTEGRAL

The **discrete Choquet integral** of $f : C = \{c_1, \ldots, c_m\} \to \mathbb{R}_+$ with respect to the capacity $\mu : 2^C \to [0, 1]$ is defined as follows:

$$\mathcal{C}_\mu(f) = \sum_{i=1}^{m} \left(f(c_{(i)}) - f(c_{(i-1)})\right) \cdot \mu\left(A_{(i)}\right) \ ,$$

where $(\cdot)$ is a permutation of $\{1, \ldots, m\}$ such that $0 \leq f(c_{(1)}) \leq f(c_{(2)}) \leq \ldots \leq f(c_{(m)})$, and $A_{(i)} = \{c_{(i)}, \ldots, c_{(m)}\}$.

*How to aggregate individual values, giving the right weight/importance to each **subset** of criteria?*

DISCRETE CHOQUET INTEGRAL

INTELLIGENT SYSTEMS

The **discrete Choquet integral** of $f : C = \{c_1, \ldots, c_m\} \to \mathbb{R}_+$ with respect to the capacity $\mu : 2^C \to [0,1]$ is defined as follows:

$$\mathcal{C}_\mu(f) = \sum_{i=1}^{m} \left(f(c_{(i)}) - f(c_{(i-1)}) \right) \cdot \mu\left(A_{(i)} \right) \ ,$$

where $(\cdot)$ is a permutation of $\{1, \ldots, m\}$ such that $0 \leq f(c_{(1)}) \leq f(c_{(2)}) \leq \ldots \leq f(c_{(m)})$, and $A_{(i)} = \{c_{(i)}, \ldots, c_{(m)}\}$.

The Choquet integral expressed in terms of the Möbius transform:

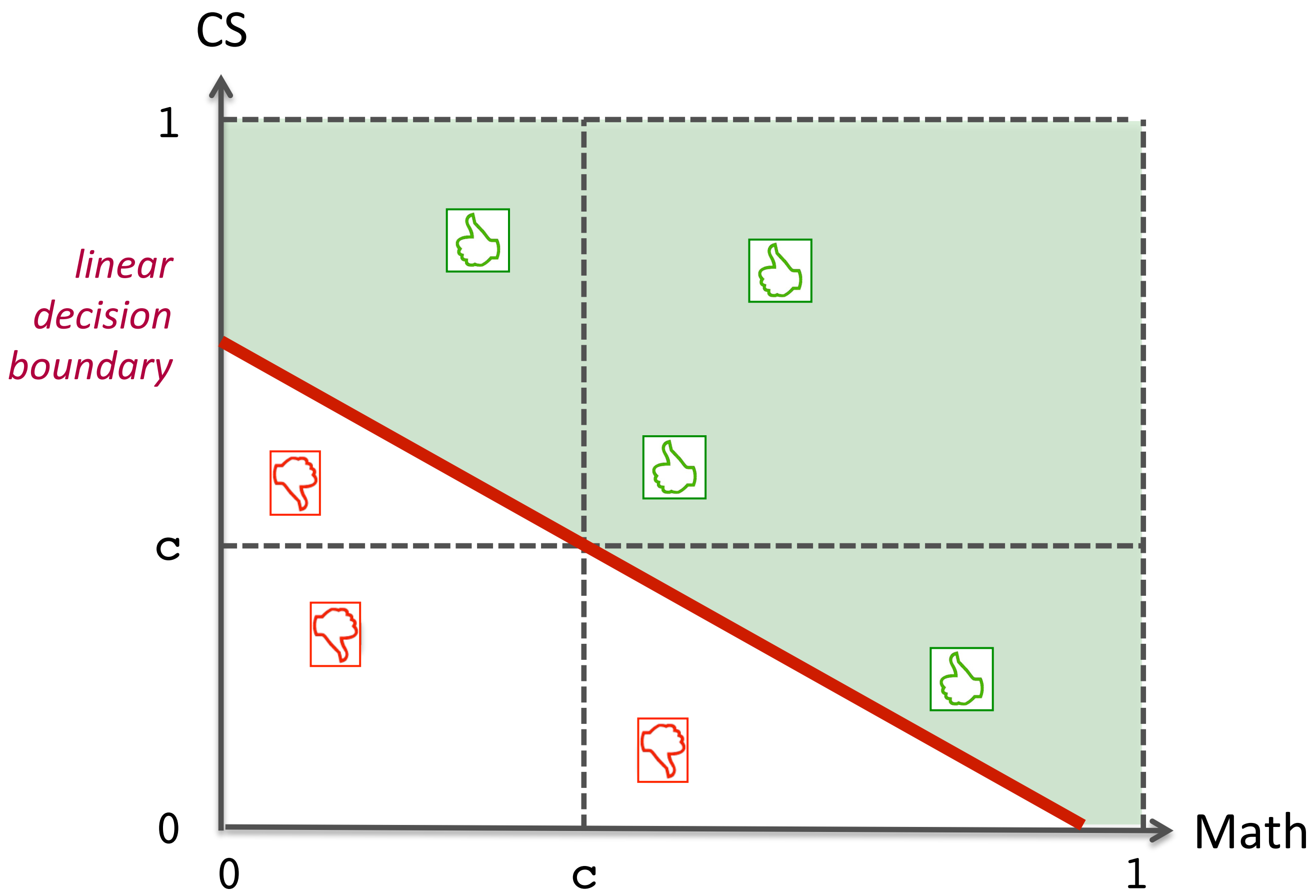
$$\mathcal{C}_\mu(f) = \sum_{T \subseteq C} \boldsymbol{m}_\mu(T) \times \min_{c_i \in T} f(c_i)$$

AGGREGATION OPERATORS

Special cases:

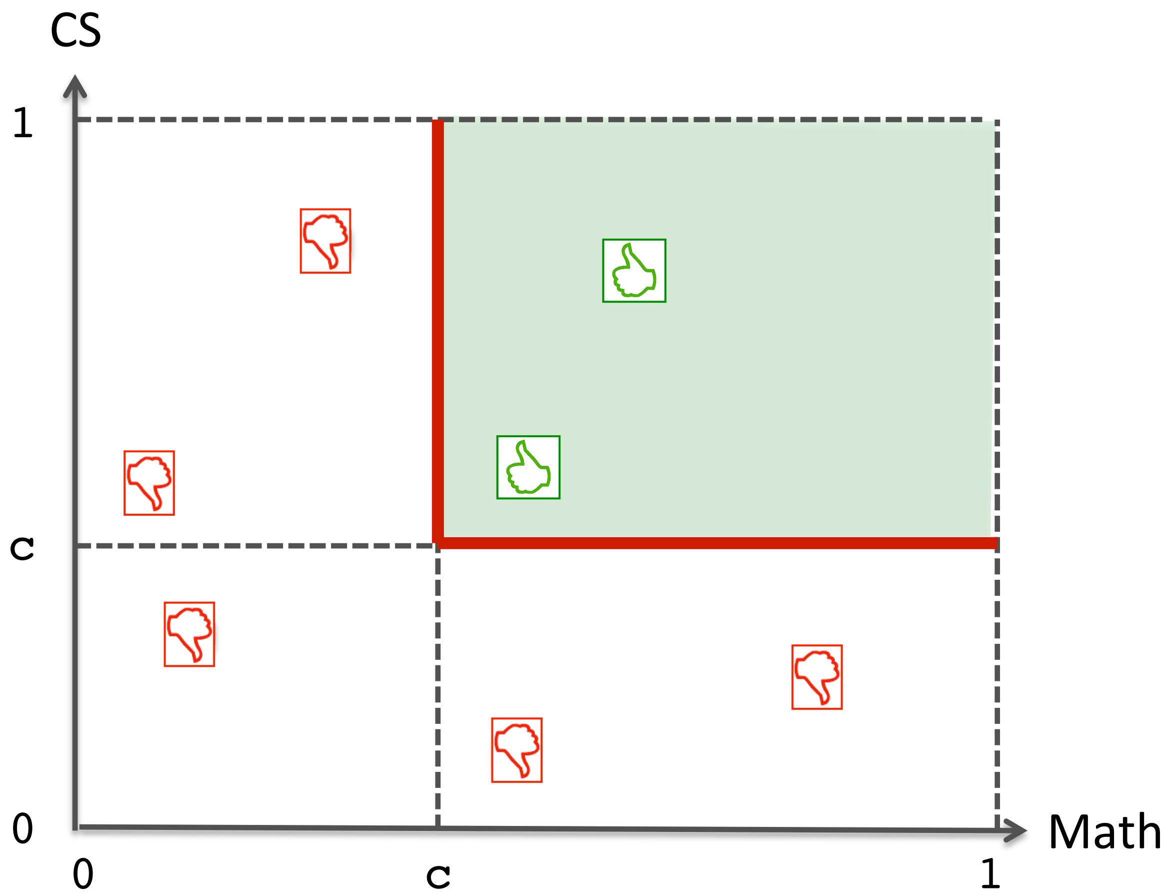
- min
- max
- weighted average (additive measure)
- ordered weighted average (OWA)

DECISION BOUNDARY IN 2D



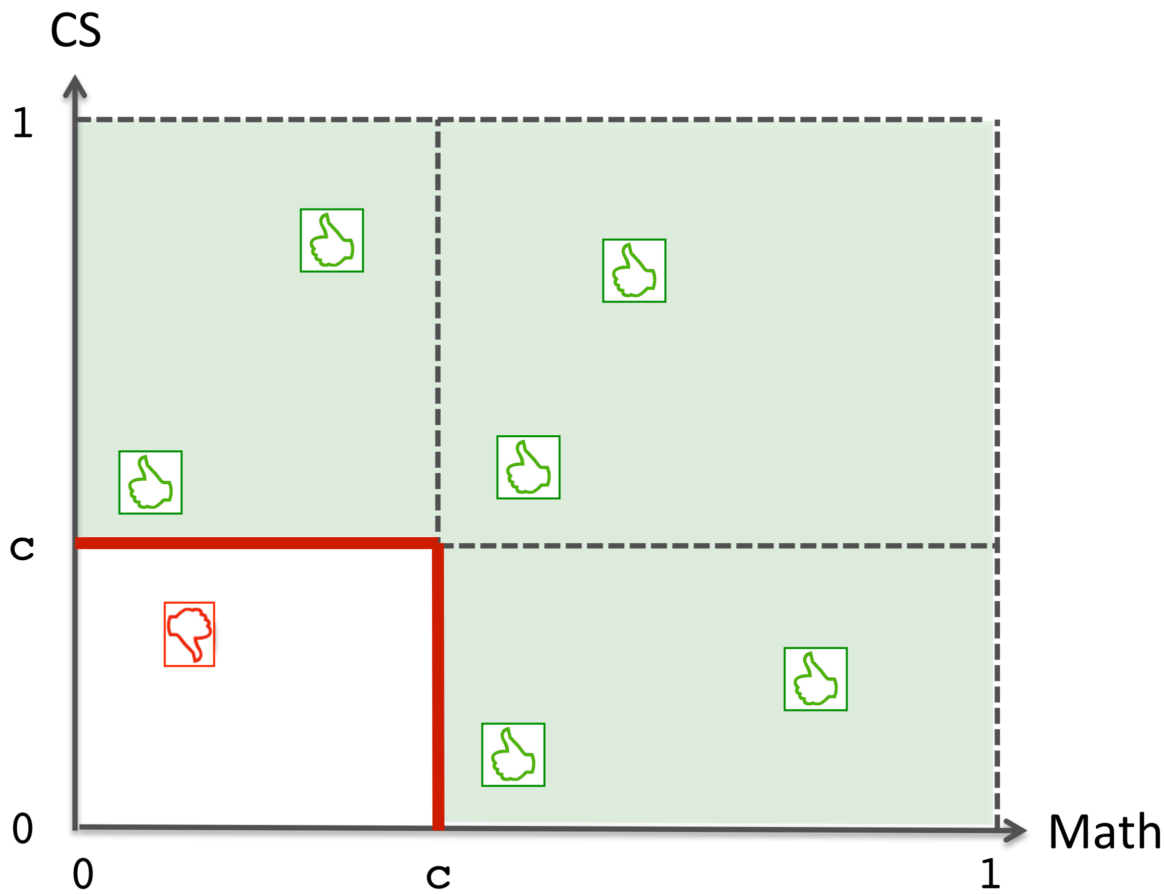
$$(x, y) \mapsto \mathbb{I}\Big(\alpha x + (1-\alpha) y > c\Big)$$

DECISION BOUNDARY IN 2D



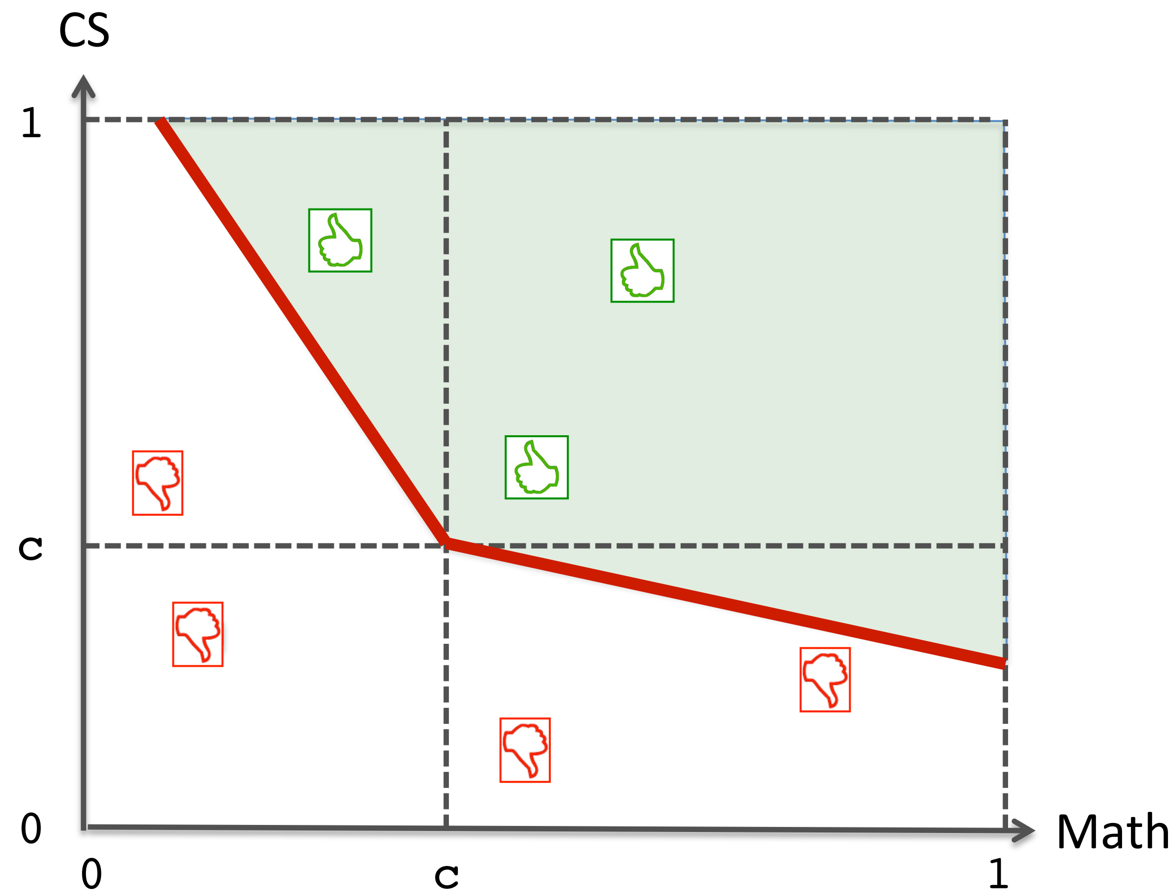
$$(x, y) \mapsto \mathbb{I}\Big(\min(x, y) > c\Big)$$

DECISION BOUNDARY IN 2D

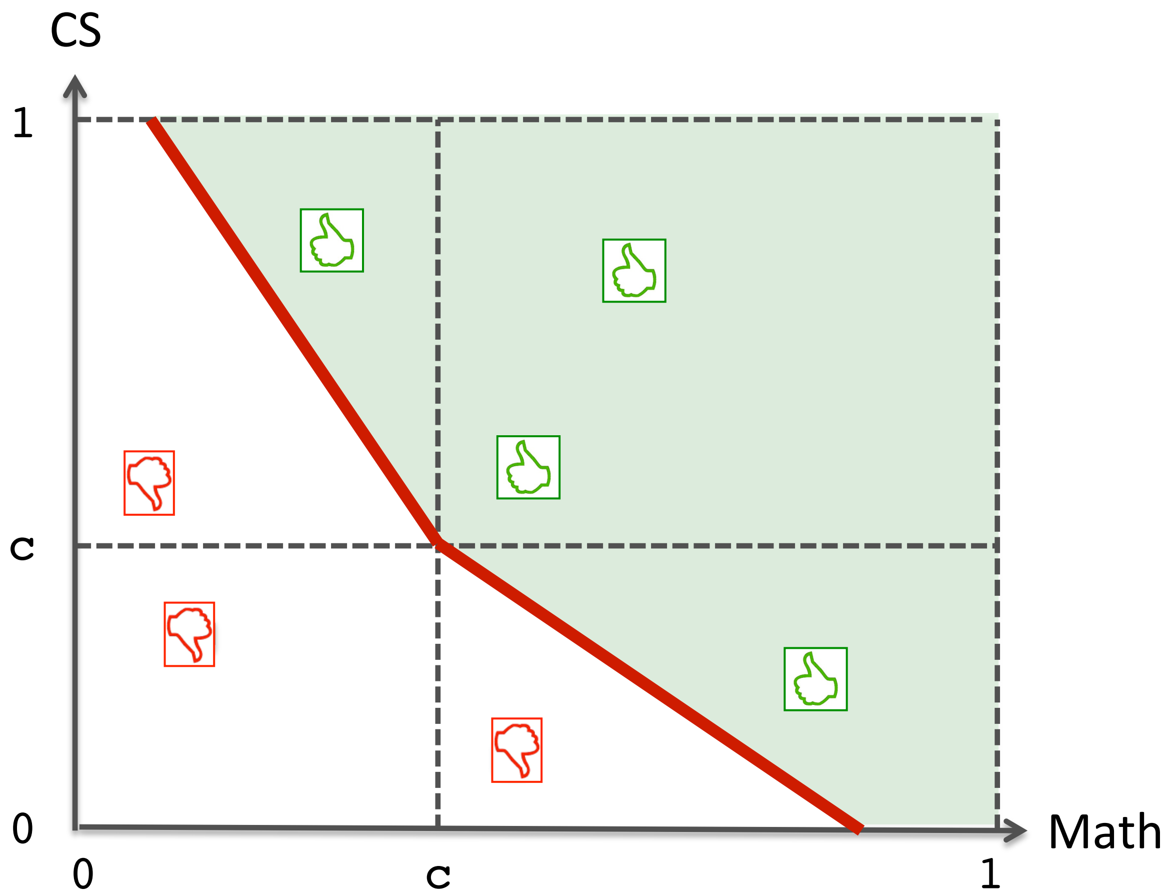


$$(x, y) \mapsto \mathbb{I}\Big(\max(x, y) > c\Big)$$

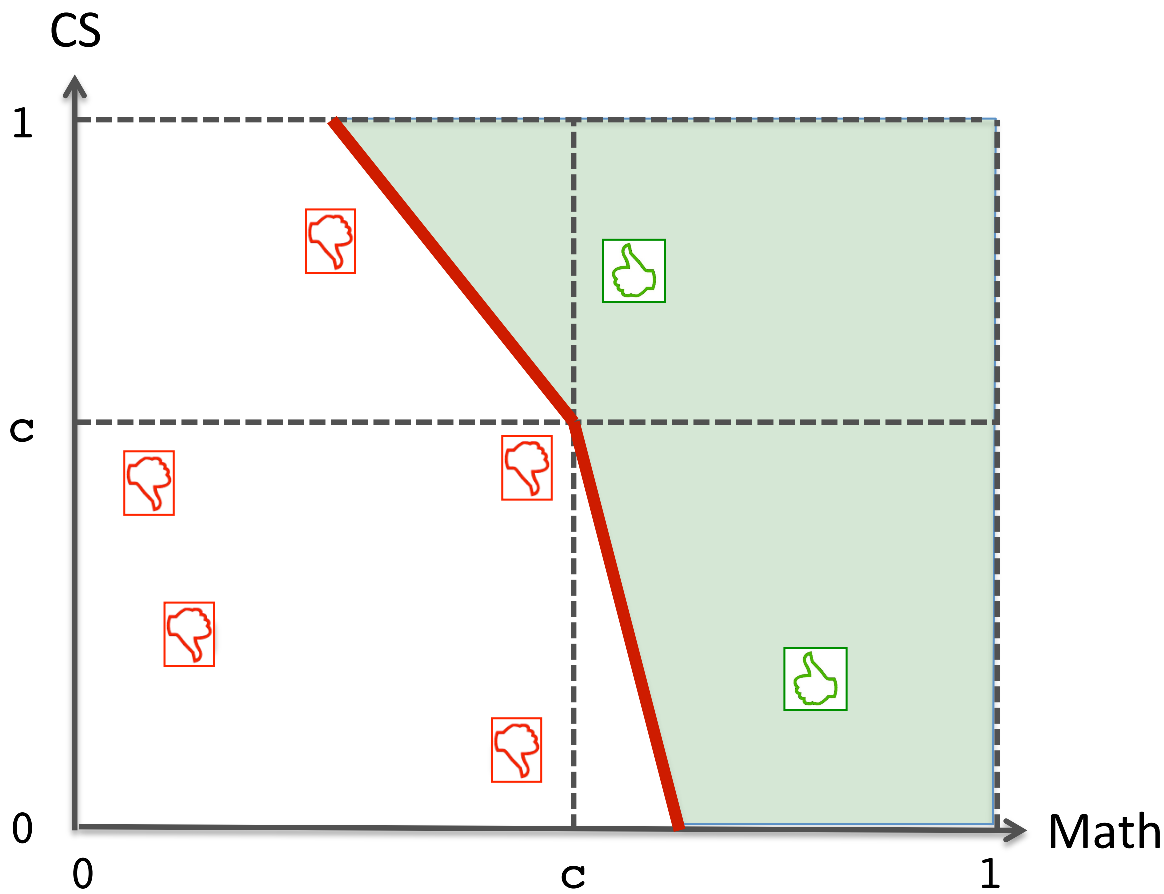
DECISION BOUNDARY IN 2D



DECISION BOUNDARY IN 2D



DECISION BOUNDARY IN 2D



VC DIMENSION

THEOREM: For the model class $\mathcal{H}$ given by the thresholded Choquet integral,

$$VC(\mathcal{H}) = \Omega\left(\frac{2^m}{\sqrt{m}}\right).$$

That is, the VC dimension of $\mathcal{H}$ grows asymptotically at least as fast as $2^m/\sqrt{m}$.

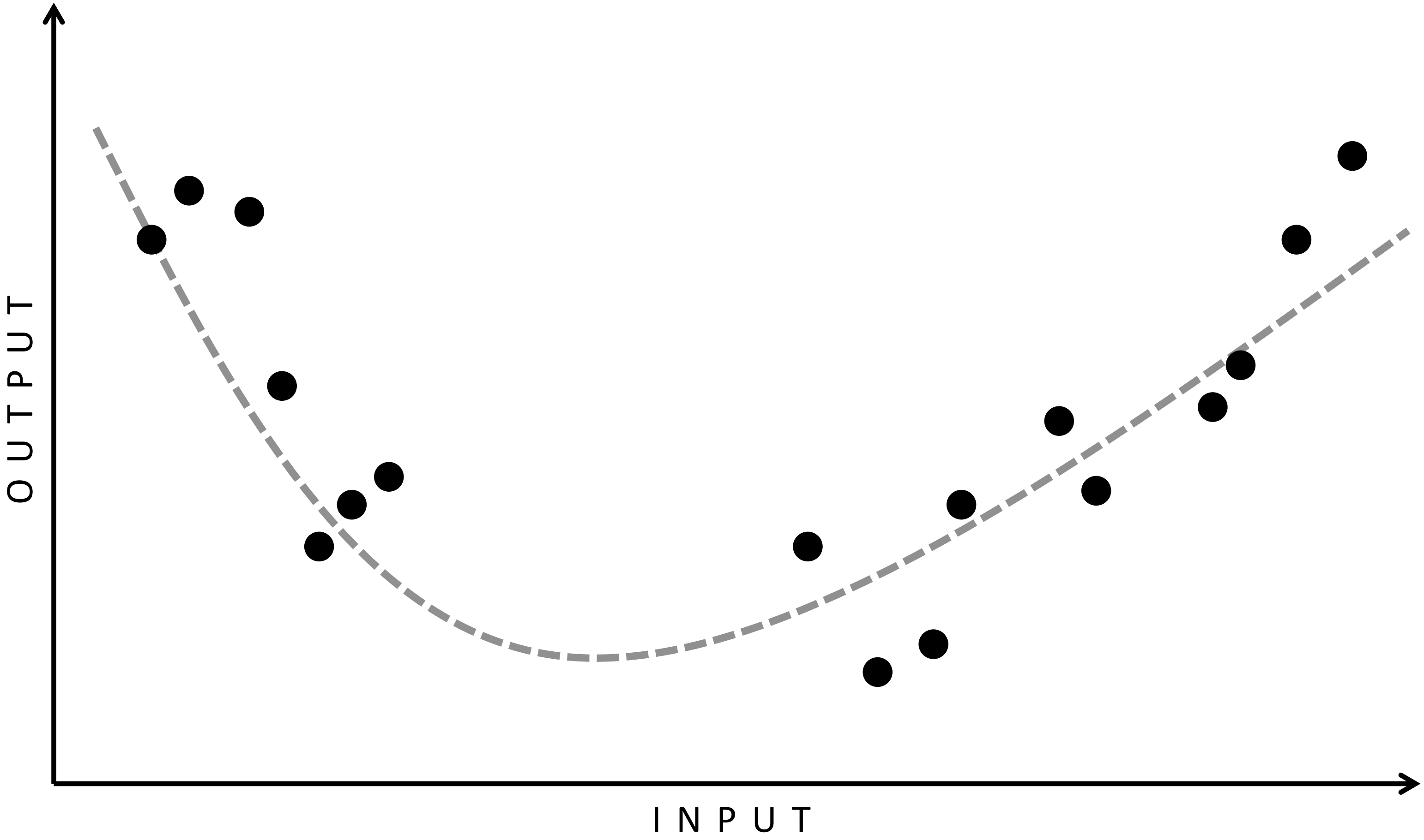
This bound is relatively tight, since $VC(\mathcal{H}) \leq 2^m$ is a trivial upper bound.

Restricting to k-additive measures for small k, we can show that

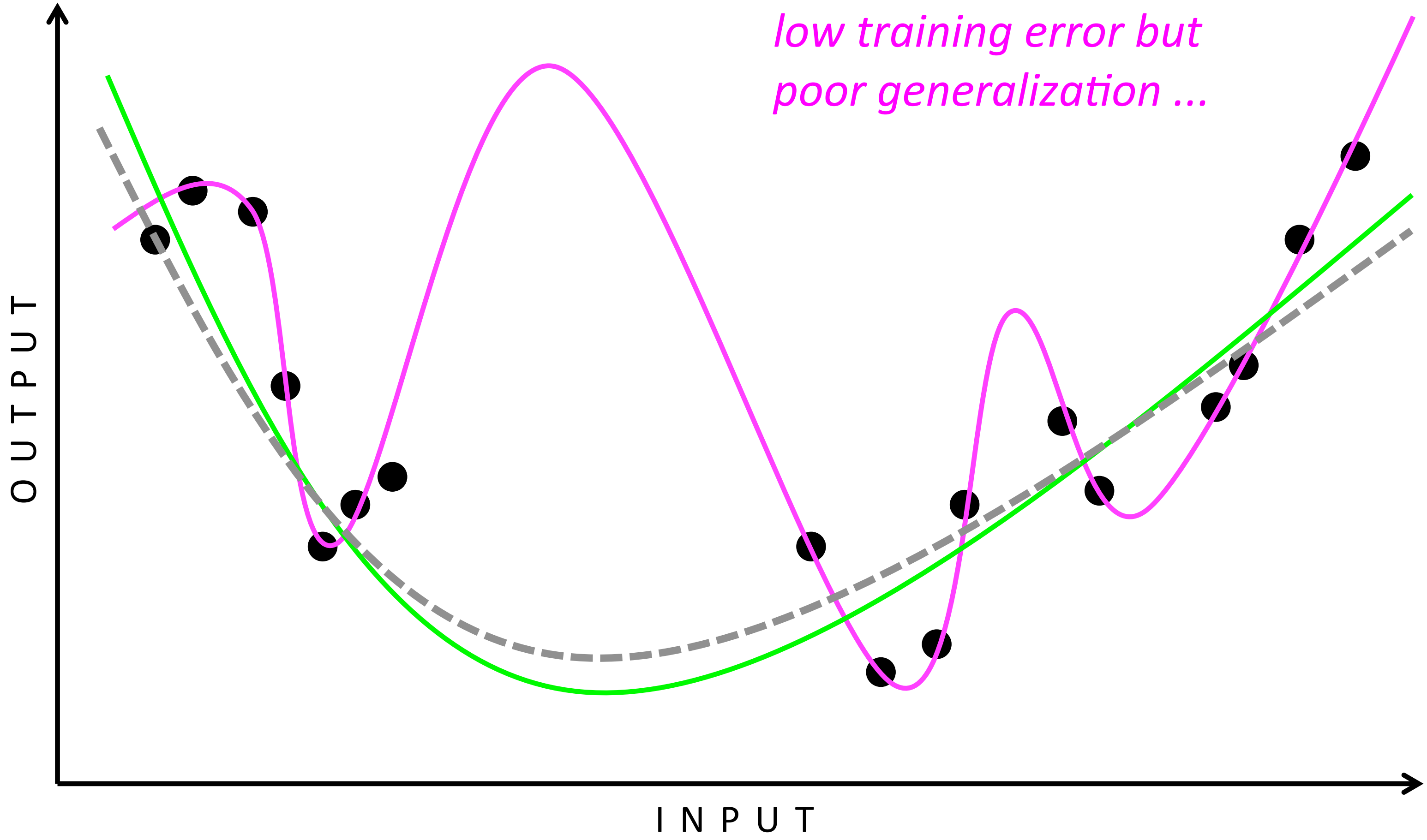
$$VC(\mathcal{H}) = \Omega\left(m^k\right).$$

$\rightarrow$ Choice of k as a means for capacity control!

GENERALIZATION



GENERALIZATION



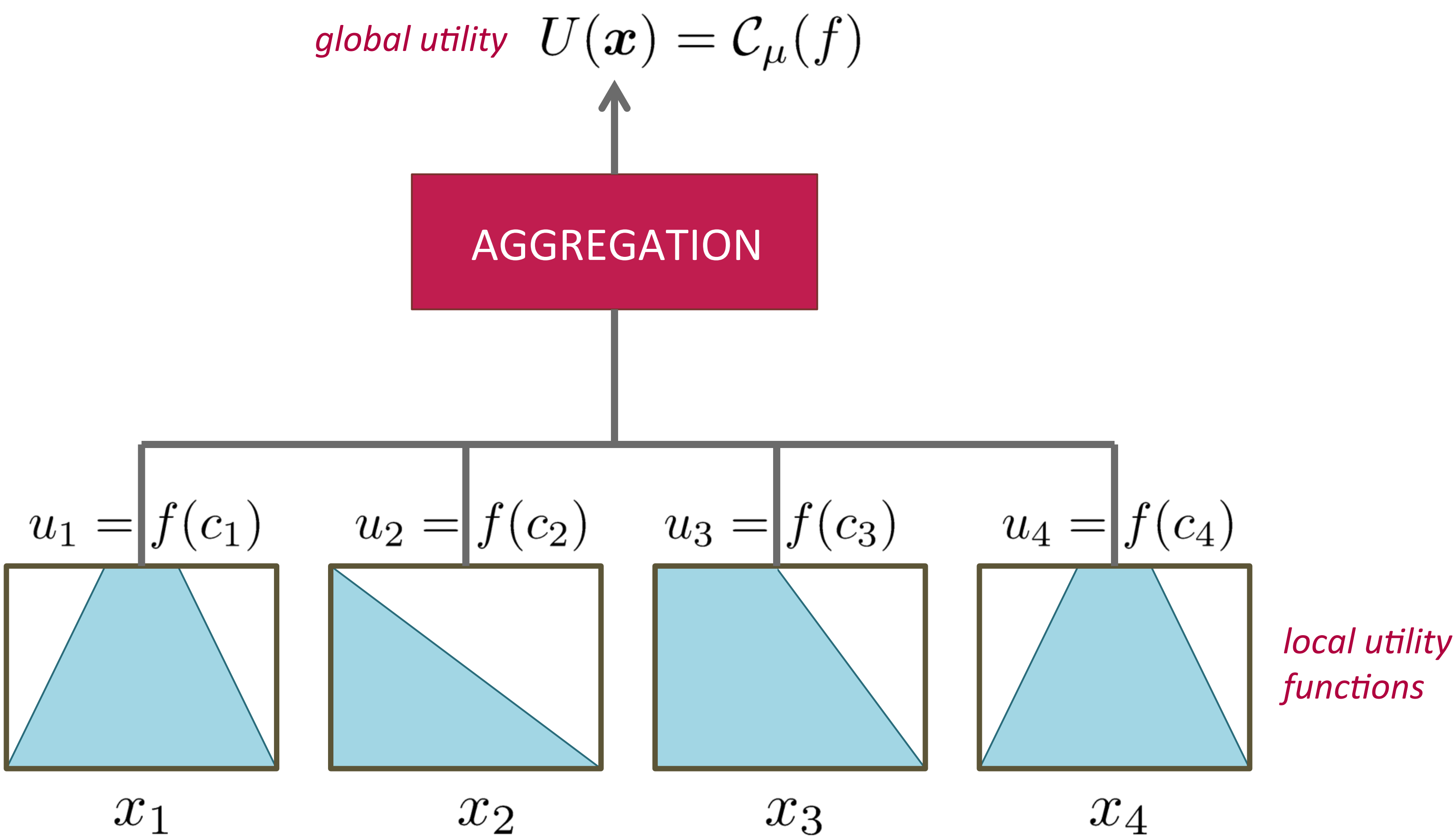
LEARNING WITH THE CHOQUET INTEGRAL

- We advocate the Choquet integral as a versatile tool in the context of (supervised) machine learning, especially for learning **monotone models**.

- Being used as a prediction function to combine several input features (criteria) into an output, the Choquet integral nicely combines

  - *monotonicity*
  - *non-linearity*
  - *interpretability (importance, interaction)*

LEARNING WITH THE CHOQUET INTEGRAL



TRAINING DATA:

$$
\begin{array}{ccc}
(10, 29, \ldots, 60) & \succ & (72, 18, \ldots, 52) \\
(40, 33, \ldots, 72) & \succ & (50, 40, \ldots, 37) \\
(60, 39, \ldots, 70) & \succ & (52, 48, \ldots, 62) \\
\cdots & \succ & \cdots
\end{array}
$$

The goal is to find a Choquet integral whose utility degrees tend to agree with the observed pairwise preferences!

ORDINAL CLASSIFICATION / SORTING

TRAINING DATA:

$$
\begin{aligned}
(10, 29, \ldots, 60) & \rightarrow & **\\
(40, 33, \ldots, 72) & \rightarrow & ***\\
(60, 39, \ldots, 70) & \rightarrow & *\\
\cdots & \rightarrow & \cdots
\end{aligned}
$$

The goal is to find a Choquet integral whose utility degrees tend to agree with the observed classifications!

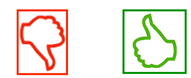
THE BINARY CASE

INTELLIGENT SYSTEMS

TRAINING DATA:

$$
\begin{array}{rcl}
(10, 29, \ldots, 60) & \rightarrow & 0 \\
(40, 33, \ldots, 72) & \rightarrow & 1 \\
(60, 39, \ldots, 70) & \rightarrow & 0 \\
\cdots & \rightarrow & \cdots
\end{array}
$$

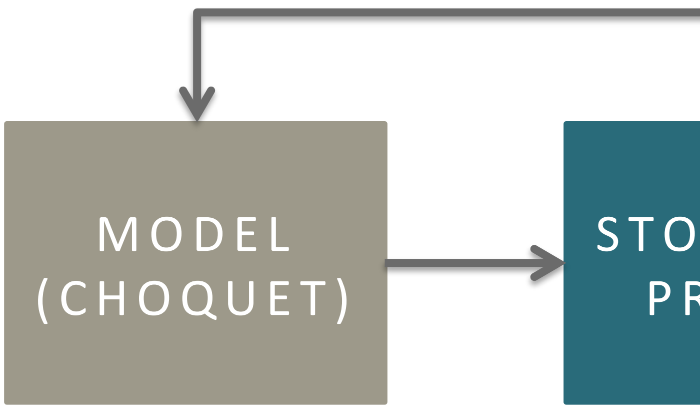
distinguishing between „good“ and „bad“



The goal is to find a Choquet integral whose utility degrees tend to agree with the observed classifications!

MODEL IDENTIFICATION

Probabilistic modeling allows for the use of induction principles such as maximum likelihood (ML) estimation!

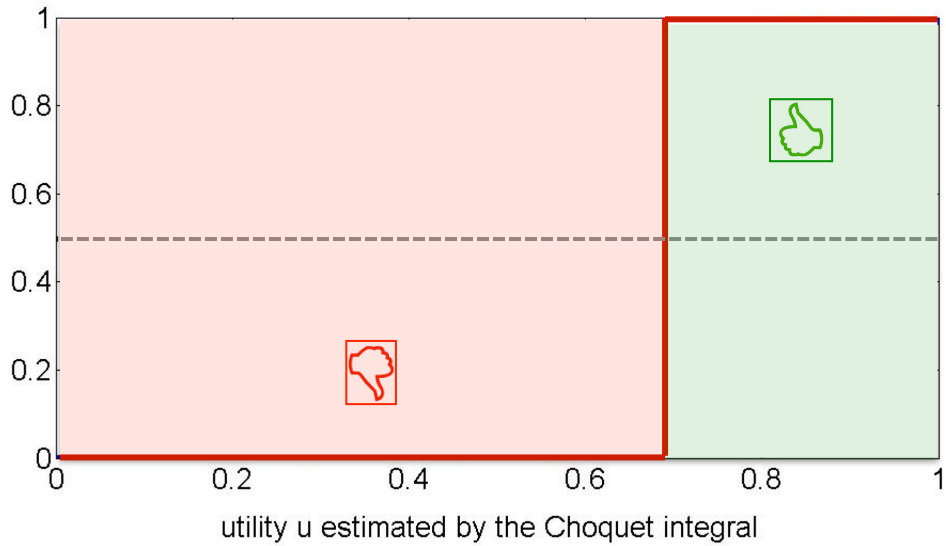


... in contrast to (linear) programming techniques as used in preference elicitation.

Decision making (discrete rating) as a **two-stage process**:

(1) a (latent) utility degree $u = \mathcal{C}_\mu(\boldsymbol{x}) \in [0,1]$ is determined by the Choquet integral

(2) a discrete choice is made by thresholding u at β



STOCHASTIC MODELING

Decision making (discrete rating) as a **two-stage process**:

(1) a (latent) utility degree $u = \mathcal{C}_\mu(\boldsymbol{x}) \in [0,1]$ is determined by the Choquet integral

(2) a discrete choice is made by **soft** thresholding u at β

LOGISTIC NOISY RESPONSE MODEL

$$\mathbf{P}(y=1) = \frac{1}{1 + \exp\left(-\gamma \left(\mathcal{C}_\mu(\boldsymbol{x}) - \beta \right) \right)}$$

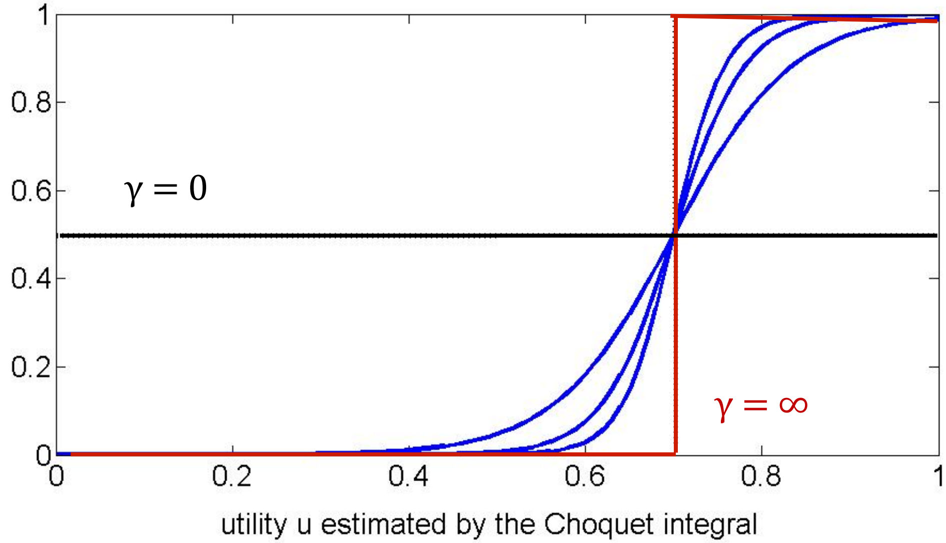
precision of the model (γ)

utility threshold (β)

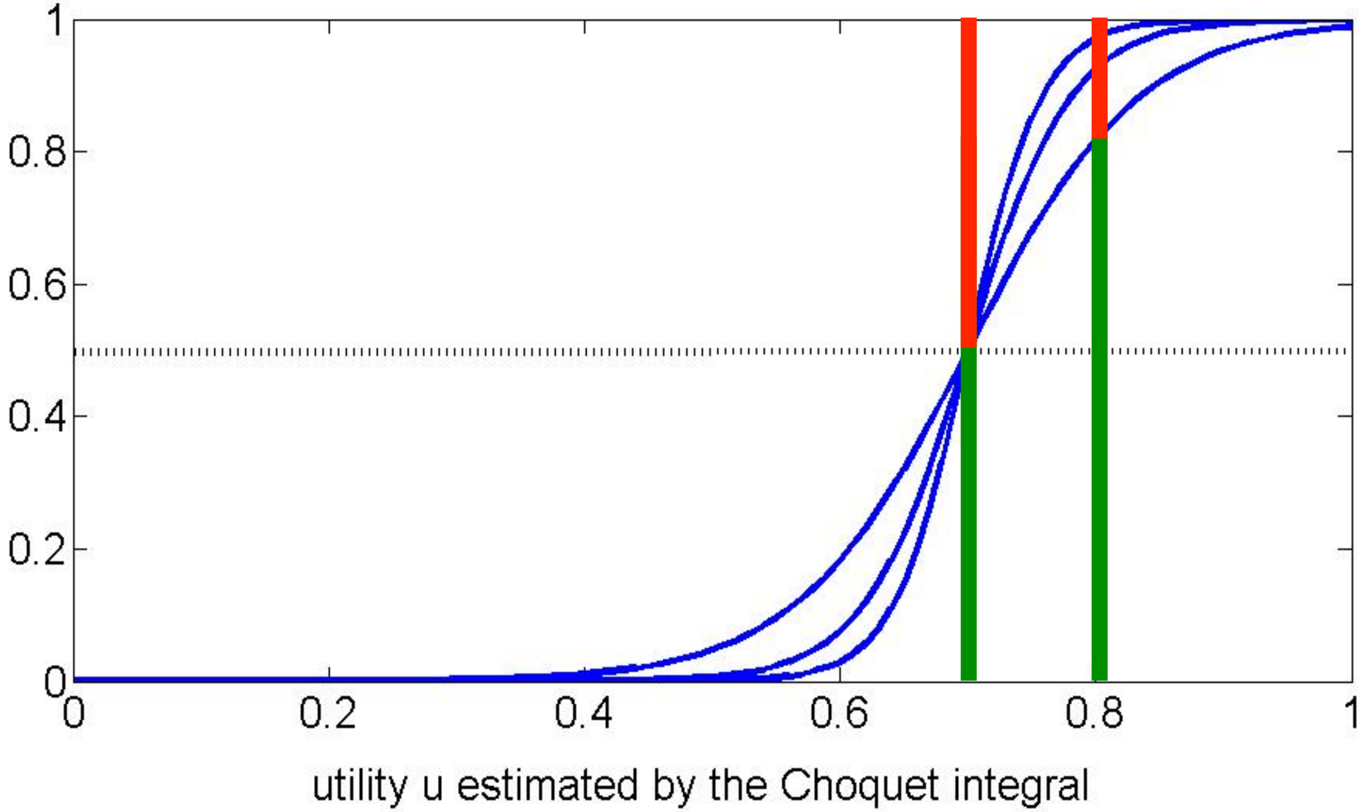
Decision making (discrete rating) as a **two-stage process**:

(1) a (latent) utility degree $u = \mathcal{C}_\mu(\boldsymbol{x}) \in [0, 1]$ is determined by the Choquet integral

(2) a discrete choice is made by **soft** thresholding u at β



STOCHASTIC MODELING



CHOQUISTIC REGRESSION

INTELLIGENT SYSTEMS

Logistic

$$\log\left(\frac{\mathbf{P}(y=1\,|\,\boldsymbol{x})}{\mathbf{P}(y=0\,|\,\boldsymbol{x})}\right) = w_0 + \boldsymbol{w}^\top \boldsymbol{x}$$

Choquistic

$$\log\left(\frac{\mathbf{P}(y=1\,|\,\boldsymbol{x})}{\mathbf{P}(y=0\,|\,\boldsymbol{x})}\right) = \gamma\left(\mathcal{C}_\mu(\boldsymbol{x}) - \beta\right)$$

Choquet integral of (normalized) attribute values

- A. Fallah Tehrani, W. Cheng, K. Dembczynski, EH. Learning Monotone Nonlinear Models using the Choquet Integral. Machine Learning, 89(1), 2012.
- A. Fallah Tehrani, W. Cheng, EH. Preference Learning using the Choquet Integral: The Case of Multipartite Ranking. IEEE Transactions on Fuzzy Systems, 2012.
- EH, A. Fallah Tehrani. Efficient Learning of Classifiers based on the 2-additive Choquet Integral. In: Computational Intelligence in Intelligent Data Analysis. Springer, 2012.
- A. Fallah Tehrani and EH. Ordinal Choquistic Regression. EUSFLAT 2013.

MAXIMUM LIKELIHOOD ESTIMATION

INTELLIGENT SYSTEMS

Given a set of (i.i.d.) training data

$$\mathcal{D} = \left\{(\boldsymbol{x}^{(i)}, y^{(i)})\right\}_{i=1}^{n} \in (\mathbb{R}^m \times \mathcal{Y})^n \quad ,$$

the likelihood of the parameters is given by

$$L(\boldsymbol{m}, \boldsymbol{\beta}, \gamma) = \prod_{i=1}^{n} \mathbf{P}\left(y = y^{(i)} \mid \boldsymbol{x}^{(i)}, \boldsymbol{m}, \boldsymbol{\beta}, \gamma\right) \quad .$$

MAXIMUM LIKELIHOOD ESTIMATION

ML estimation leads to a **constrained optimization problem**:

$$\min_{\boldsymbol{m},\gamma,\beta} \; \gamma \sum_{i=1}^{n} (1 - y^{(i)}) \left(\mathcal{C}_{\boldsymbol{m}}(\boldsymbol{x}^{(i)}) - \beta\right) + \sum_{i=1}^{n} \log\left(1 + \exp(-\gamma\left(\mathcal{C}_{\boldsymbol{m}}(\boldsymbol{x}^{(i)}) - \beta\right))\right)$$

subject to:

$$0 \leq \beta \leq 1$$

$$0 < \gamma$$

conditions on bias and scale parameter

$$\sum_{T \subseteq C} \boldsymbol{m}(T) = 1$$

$$\sum_{B \subseteq A \setminus \{c_i\}} \boldsymbol{m}(B \cup \{c_i\}) \geq 0 \quad \forall A \subseteq C, \forall c_i \in C$$

normalization and monotonicity of the non-additive measure

→ *computationally expensive!*

EXPERIMENTAL EVALUATION

	dataset	CR 🟩🟦	LR 🟩	KLR-ply 🟦	KLR-rbf 🟦	MORE 🟩🟦
20%	DBS	.2226±.0380 (4)	.1803±.0336 (1)	.2067±.0447 (3)	.1922±.0501 (2)	.2541±.0142 (5)
	CPU	.0457±.0338 (2)	.0430±.0318 (1)	.0586±.0203 (3)	.0674±.0276 (4)	.1033±.0681 (5)
	BCC	.2939±.0100 (4)	.2761±.0265 (1)	.3102±.0386 (5)	.2859±.0329 (3)	.2781±.0219 (2)
	MPG	.0688±.0098 (2)	.0664±.0162 (1)	.0729±.0116 (4)	.0705±.0122 (3)	.0800±.0198 (5)
	ESL	.0764±.0291 (3)	.0747±.0243 (1)	.0752±.0117 (2)	.0794±.0134 (4)	.1035±.0332 (5)
	MMG	.1816±.0140 (3)	.1752±.0106 (2)	.1970±.0095 (4)	.2011±.0123 (5)	.1670±.0120 (1)
	ERA	.2997±.0123 (2)	.2922±.0096 (1)	.3011±.0132 (3)	.3259±.0172 (5)	.3040±.0192 (4)
	LEV	.1527±.0138 (1)	.1644±.0106 (4)	.1570±.0116 (2)	.1577±.0124 (3)	.1878±.0242 (5)
	CEV	.0441±.0128 (1)	.1689±.0066 (5)	.0571±.0078 (3)	.0522±.0085 (2)	.0690±.0408 (4)
	avg. rank	**2.4**	**1.9**	**3.3**	**3.4**	**4**
50%	DBS	.1560±.0405 (3)	.1443±.0371 (2)	.1845±.0347 (5)	.1628±.0269 (4)	.1358±.0432 (1)
	CPU	.0156±.0135 (1)	.0400±.0106 (3)	.0377±.0153 (2)	.0442±.0223 (5)	.0417±.0198 (4)
	BCC	.2871±.0358 (4)	.2647±.0267 (2)	.2706±.0295 (3)	.2879±.0269 (5)	.2616±.0320 (1)
	MPG	.0641±.0175 (1)	.0684±.0206 (2)	.1462±.0218 (5)	.1361±.0197 (4)	.0700±.0162 (3)
	ESL	.0660±.0135 (1)	.0697±.0144 (3)	.0704±.0128 (5)	.0699±.0148 (4)	.0690±.0171 (2)
	MMG	.1736±.0157 (3)	.1710±.0161 (2)	.1859±.0141 (4)	.1900±.0169 (5)	.1604±.0139 (1)
	ERA	.3008±.0135 (3)	.3054±.0140 (4)	.2907±.0136 (1)	.3084±.0152 (5)	.2928±.0168 (2)
	LEV	.1357±.0122 (1)	.1641±.0131 (4)	.1500±.0098 (3)	.1482±.0112 (2)	.1658±.0202 (5)
	CEV	.0346±.0076 (1)	.1667±.0093 (5)	.0357±.0113 (2)	.0393±.0090 (3)	.0443±.0080 (4)
	avg. rank	**2**	**3**	**3.3**	**4.1**	**2.6**
80%	DBS	.1363±.0380 (2)	.1409±.0336 (4)	.1422±.0498 (5)	.1386±.0521 (3)	.0974±.0560 (1)
	CPU	.0089±.0126 (1)	.0366±.0068 (4)	.0329±.0295 (2)	.0384±.0326 (5)	.0342±.0232 (3)
	BCC	.2631±.0424 (2)	.2669±.0483 (3)	.2784±.0277 (4)	.2937±.0297 (5)	.2526±.0472 (1)
	MPG	.0526±.0263 (1)	.0538±.0282 (2)	.0669±.0251 (4)	.0814±.0309 (5)	.0656±.0248 (3)
	ESL	.0517±.0235 (1)	.0602±.0264 (2)	.0654±.0228 (3)	.0718±.0188 (5)	.0657±.0251 (4)
	MMG	.1584±.0255 (2)	.1683±.0231 (3)	.1798±.0293 (4)	.1853±.0232 (5)	.1521±.0249 (1)
	ERA	.2855±.0257 (1)	.2932±.0261 (4)	.2885±.0302 (2)	.2951±.0286 (5)	.2894±.0278 (3)
	LEV	.1312±.0186 (1)	.1662±.0171 (5)	.1518±.0104 (3)	.1390±.0129 (2)	.1562±.0252 (4)
	CEV	.0221±.0091 (1)	.1643±.0184 (5)	.0376±.0091 (3)	.0262±.0067 (2)	.0408±.0090 (4)
	avg. rank	**1.3**	**3.6**	**3.3**	**4.1**	**2.7**

🟩 monotone classifier 🟦 nonlinear classifier

OUTLINE

INTELLIGENT SYSTEMS

PART 1 Introduction to preference learning	PART 2 Machine learning vs. MCDA
PART 3 Multi-criteria preference learning	PART 4 Preference-based online learning

MULTI-ARMED BANDITS

„pulling an arm“ ⟷ choosing an option

partial information online learning
sequential decision process

MULTI-ARMED BANDITS

INTELLIGENT SYSTEMS

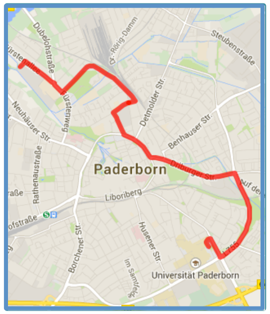
„pulling an arm" ⟷ putting an advertisement on a website

choice of an option/strategy (arm) yields a **random reward**

partial information online learning
sequential decision process

MULTI-ARMED BANDITS

INTELLIGENT
SYSTEMS



„pulling an arm“ ⟷ picking a traffic route from source to target

choice of an option/strategy (arm) yields a **random reward**

partial information online learning
sequential decision process

MULTI-ARMED BANDITS

INTELLIGENT SYSTEMS

$X_1 \sim \mathbf{P}_1$	$X_2 \sim \mathbf{P}_2$	$X_3 \sim \mathbf{P}_3$	$X_4 \sim \mathbf{P}_4$	$X_5 \sim \mathbf{P}_5$

„pulling an arm“ $\longleftrightarrow$ choosing an option

choice of an option/strategy (arm) yields a **random reward**

partial information online learning
sequential decision process

MULTI-ARMED BANDITS

$X_1 \sim \mathbf{P}_1$ | $X_2 \sim \mathbf{P}_2$ | $X_3 \sim \mathbf{P}_3$ | $X_4 \sim \mathbf{P}_4$ | $X_5 \sim \mathbf{P}_5$

Immediate reward: 2.5

Cumulative reward: 2.5

MULTI-ARMED BANDITS

INTELLIGENT SYSTEMS

$X_1 \sim \mathbf{P}_1$ $X_2 \sim \mathbf{P}_2$ $X_3 \sim \mathbf{P}_3$ $X_4 \sim \mathbf{P}_4$ $X_5 \sim \mathbf{P}_5$

Immediate reward: `2.5 3.1`

Cumulative reward: `2.5 5.6`

MULTI-ARMED BANDITS

$$X_1 \sim \mathbf{P}_1 \quad X_2 \sim \mathbf{P}_2 \quad X_3 \sim \mathbf{P}_3 \quad X_4 \sim \mathbf{P}_4 \quad X_5 \sim \mathbf{P}_5$$

Immediate reward:	2.5	3.1	1.7
Cumulative reward:	2.5	5.6	7.3

MULTI-ARMED BANDITS

INTELLIGENT SYSTEMS

$X_1 \sim \mathbf{P}_1$ $X_2 \sim \mathbf{P}_2$ $X_3 \sim \mathbf{P}_3$ $X_4 \sim \mathbf{P}_4$ $X_5 \sim \mathbf{P}_5$

Immediate reward:	2.5	3.1	1.7	3.7	...
Cumulative reward:	2.5	5.6	7.3	11.0	...

maximize cumulative reward → *explore and exploit (tradeoff)*

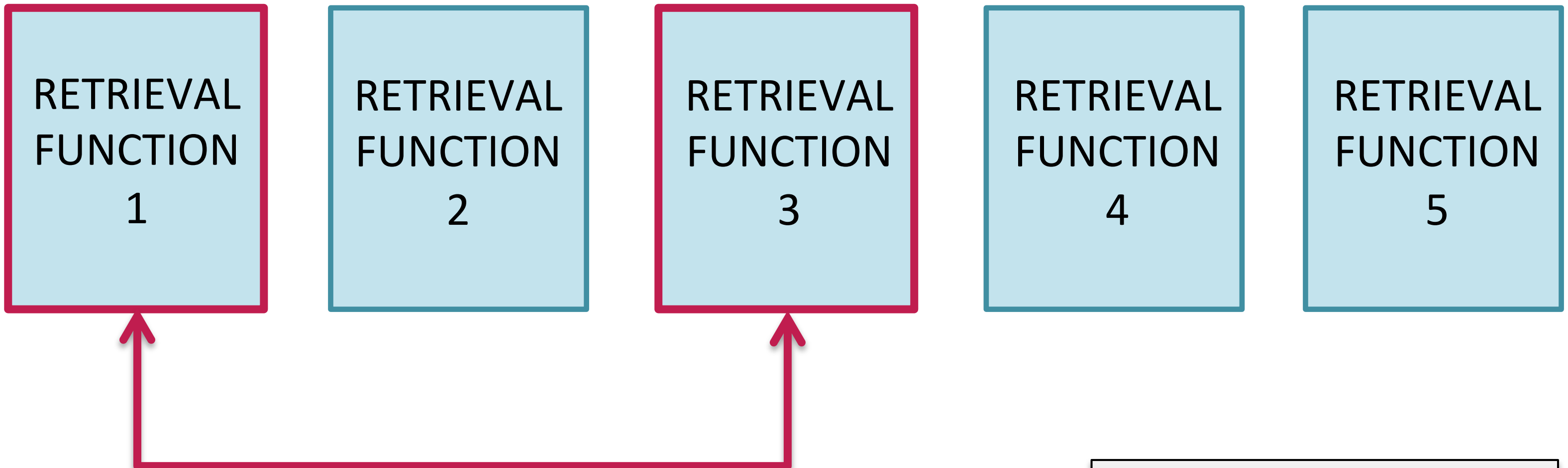
find best option → *pure exploration (effort vs. certainty)*

PREFERENCE-BASED BANDITS

$X_1 \sim \mathbf{P}_1$	$X_2 \sim \mathbf{P}_2$	$X_3 \sim \mathbf{P}_3$	$X_4 \sim \mathbf{P}_4$	$X_5 \sim \mathbf{P}_5$

In many applications,

- the assignment of (numeric) **rewards to single outcomes** (and hence the assessment of individual options on an absolute scale) is difficult,
- while the **qualitative comparison between pairs of outcomes** (arms/ options) is more feasible.

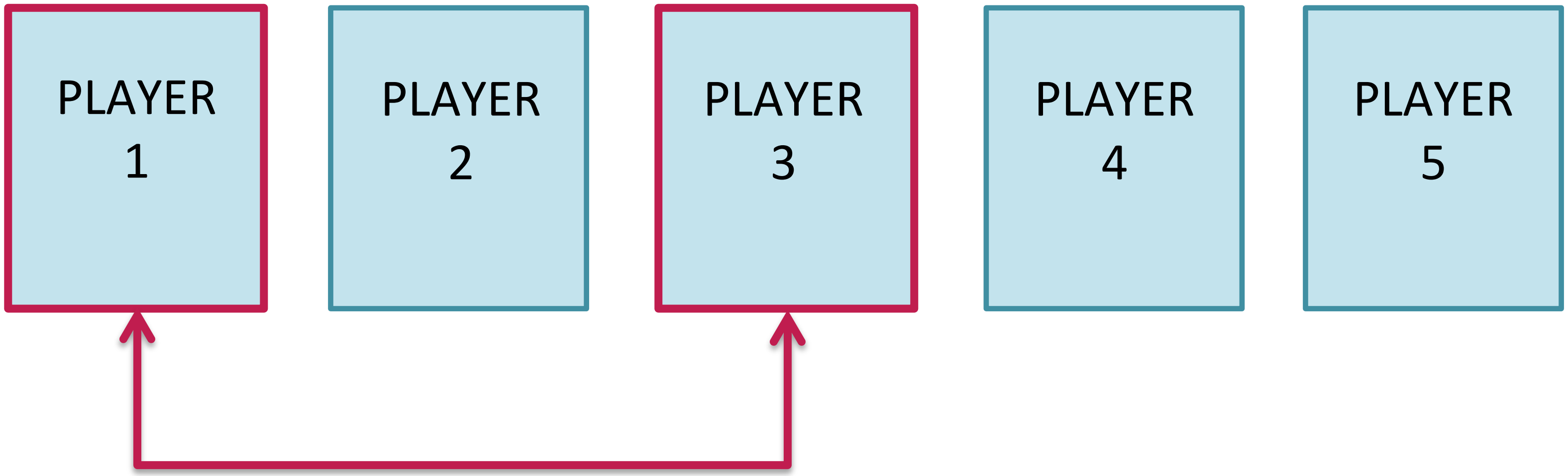


$$X_3 \succ X_1$$

The result returned by the third retrieval function, for a given query, is preferred to the result returned by the first search engine.

Noisy preferences can be inferred from how a user clicks through an **interleaved** list of documents [Radlinski et al., 2008].

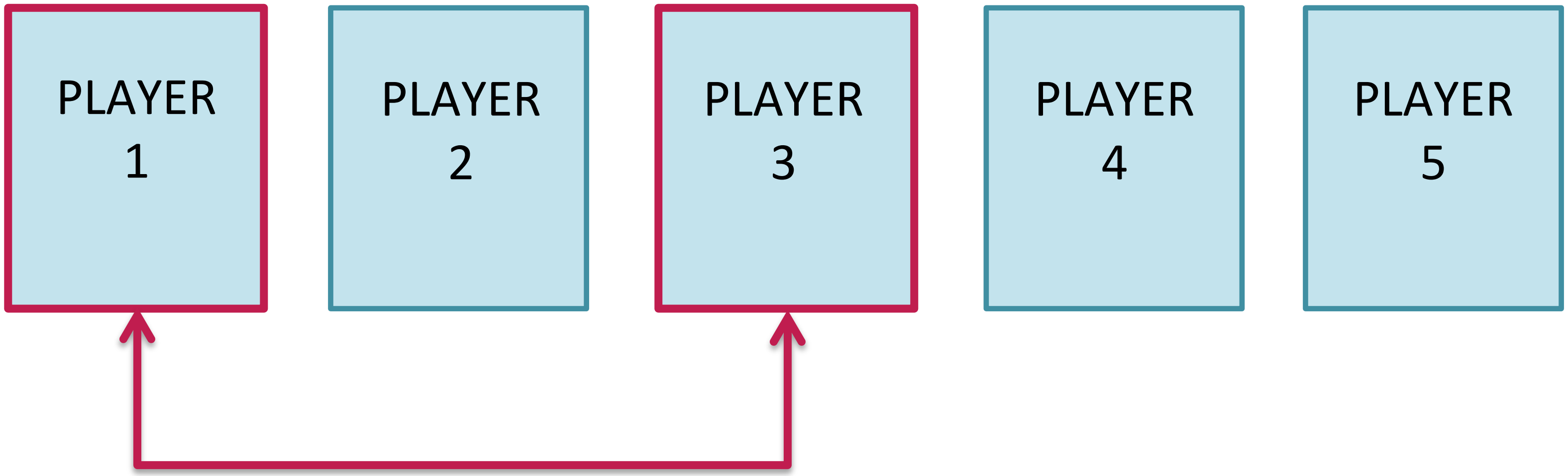
PREFERENCE-BASED BANDITS



$$X_3 \succ X_1$$

Third player has beaten first player in a match.

PREFERENCE-BASED BANDITS



$$X_3 \succ X_1$$

- *This setting has first been introduced as the **dueling bandits problem** (Yue and Joachims, 2009).*
- *More generally, we speak of **preference-based multi-armed bandits** (PB-MAB).*

PREFERENCE-BASED RANK ELICITATION

INTELLIGENT SYSTEMS

- Busa-Fekete et al. (2014) consider the problem of **predicting a ranking of all arms** in the pure exploration setting.
- They propose a sampling strategy called **MallowsMPR**, which is based on the **merge sort** algorithm for selecting the arms to be compared.
- However, two arms a_i and a_j are not only compared once, but possibly several times until being sure enough.
- Confidence intervals are derived from the **Hoeffding inequality**.
- Pairwise probablities $p_{i,j}$ are supposed to be the marginals of a **Mallows distribution** on rankings/permutations.

Theorem: For any $0 < \delta < 1$, MallowsMPR outputs the target ranking with probability at least $1 - \delta$, and the number of pairwise comparisons taken by the algorithm is

$$\mathcal{O}\left(\frac{K \log_2 K}{\rho^2} \log \frac{K \log_2 K}{\delta \rho}\right) ,$$

where $K =$ number of arms and $\rho = \frac{1-\exp(-\theta)}{1+\exp(-\theta)}$, with θ the concentration parameter of the Mallows distribution.

SUMMARY & CONCLUSION

Preference learning is

- **methodologically** interesting,
- **theoretically** challenging,
- and **practically** useful, with many potential **applications**;
- **interdisciplinary** (connections to operations research, decision sciences, economics, social choice, recommender systems, information retrieval, ...).

Established methods exist, but the field is still developing (e.g., online preference learning, preference-based reinforcement learning, ...)

In particular, there are many links between preference learning and decision analysis, most of which are still to be explored!

SELECTED REFERENCES

INTELLIGENT SYSTEMS

- A. Fallah Tehrani, W. Cheng, EH. Choquistic Regression: Generalizing Logistic Regression using the Choquet Integral. EUSFLAT 2011.
- A. Fallah Tehrani, W. Cheng, K. Dembczynski, E. Hüllermeier. Learning Monotone Nonlinear Models using the Choquet Integral. Machine Learning, 89(1):183—211, 2012.
- A. Fallah Tehrani, W. Cheng, K. Dembczynski, EH. Learning Monotone Nonlinear Models using the Choquet Integral. ECML/PKDD 2011.
- A. Fallah Tehrani, W. Cheng, E. Hüllermeier. Preference Learning using the Choquet Integral: The Case of Multipartite Ranking. IEEE Transactions on Fuzzy Systems, 2012.
- E. Hüllermeier, A. Fallah Tehrani. On the VC Dimension of the Choquet Integral. IPMU 2012.

- R. Busa-Fekete and E. Hüllermeier. A Survey of Preference-based Online Learning with Bandit Algorithms. Proc. ALT-2014, Int. Conf. Algorithmic Learning Theory, Bled, 2014.
- R. Busa-Fekete, E. Hüllermeier, and B. Szorenyi. Preference-based rank elicitation using statistical models: The case of Mallows. ICML 2014.
- R. Busa-Fekete, B. Szorenyi, and E. Hüllermeier. PAC rank elicitation through adaptive sampling of stochastic pairwise preferences. AAAI 2014.