Why did we start the M-PREF workshop series?

Ulrich Junker

A personal flashback as M-PREF 2023 closing remarks: what have been the original motivations for the M-PREF workshop series and do they remain valid in the age of ML?

Discovery of preferences for AI

No preferences in early AI

Early AI problem solvers explored a search space and made multiple choices.

They used heuristics to explore promising choices first, but not explicit preference models.

Already before AI, preference models have been studied in economics, psychology, and philosophy as a means to explain and predict human choice behavior.

So why not using them in a model of intelligence, which covers problem solving and decision making?

Decision Making	**Problem Solving**
Small set of known alternatives	Large set of unknown alternatives
Single choice	Multiple choices (one per sub-goal)
Comparing alternatives is the difficulty	Finding an alternative is the difficulty
Preferences determine best solution	Goals determine good enough solution

Goal-based AI turned into knowledge-based AI

This led to the discovery of several problems when modeling commonsense & legal reasoning.

It turned out that even basic human knowledge is logically inconsistent and only valid by default. Humans are able to deal with these inconsistencies and know which defaults to retract:

- **Taxonomic reasoning**: more specific defaults (Tweety can't fly since it is a penguin) override more general defaults (Tweety can fly since it is a bird).
- **Temporal reasoning**: laws of inertia ("frame axioms") for earlier actions have priority over laws of inertia for later actions.
- **Fault diagnosis**: correctness assumptions are preferred to fault assumptions.
- **Legal reasoning**: US laws with higher authority have precedence over US laws with lower authority, even if the latter are more recent.

Commonsense and legal reasoning need preferences to eliminate meaningless interpretations of knowledge in logical form and to specify which defaults to retract!

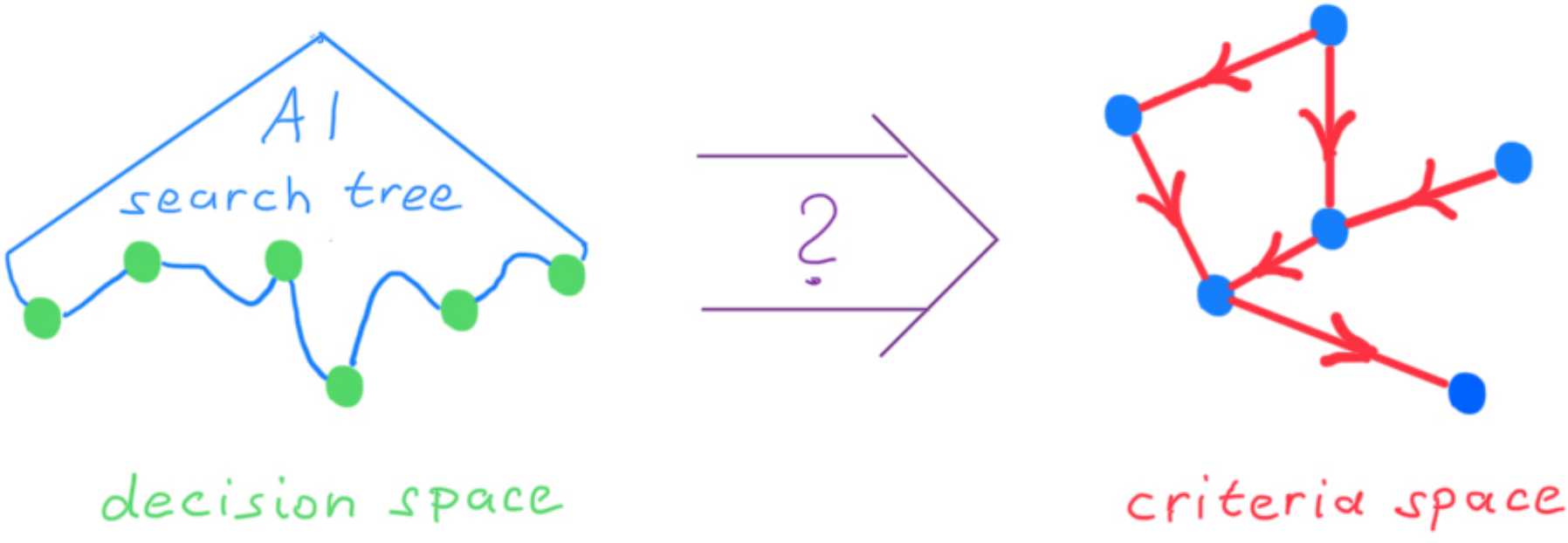
Second AI winter shifts from inference to solution finding

The focus now is on choosing a solution of a problem or a model of a theory. This is much closer to rational decision making than the original AI. But the alternatives may be complex and difficult to compute as illustrated by these examples:

- **Prioritized defaults**: choose a set of non-conflicting default rules according to a lexicographical preference order.
- **Constraint optimization**: choose a solution of hard constraints that maximizes some objective.
- **Planning and Scheduling**: choose a shortest plan or schedule.
- **Markov decision processes**: choose an optimal policy.
- **Machine learning**: choose a classifier that minimizes loss.

All these formalisms use some kind of preference order among solutions to specify which solutions in large solution spaces are interesting.

Challenges for preferences in AI



General framework for decision making

The following framework can easily be applied to AI problem solving tasks by choosing the solution space of the problem as set of alternatives. For this reason, the alternatives may be complex objects such as CSP solutions, models of logical theories, or MDP policies:

- **Alternatives**: the decision space $\mathcal{A}$ contains the alternatives. It may be infinite and defined in intensional form. It is usually difficult to compare the alternatives directly.
- **Criteria**: one or more criteria $z_i : \mathcal{A} \to \mathcal{Z}_i$ are mapping the decision space $\mathcal{A}$ to a criteria space $\mathcal{Z}_1 \times \ldots \times \mathcal{Z}_n$ and allow for a meaningful comparison of alternatives in terms of their criteria values. The mapping may be deterministic or probabilistic.
- **Preferences**: one or more (partial) preference relations $\succsim_1, \ldots, \succsim_k$ over the possibly combinatorial criteria space $\mathcal{Z}$ capture the preferences of the decision makers.

Decision-making problem: finding one or all **best alternatives**. This requires preference aggregation in presence of multiple preference relation.

New problems, new challenges!

The new kinds of decision-making problems considered in AI and other computational fields challenge some assumptions made in decision analysis:

- build the utility function first, then optimize.
- choice of a single utility function.
- qualitative preferences only in elicitation and not in reasoning.
- preferential independence of criteria.
- no importance preferences between criteria.
- limitation to positive preference statements: no negation, disjunction, quantification.
- limited tasks: elicitation, but no learning, explanation, relaxation, and revision of preferences. Reasoning with, but not about preferences.
- fixed number of criteria.
- strict ceteris paribus semantics.

New problems, new potential!

Preferences are another form of knowledge and can be subject of representation and reasoning:

Preference representations

- leverage existing AI formalism (logic, constraints, graphical models).
- capture more structure (e.g. importance preferences).

Reasoning with preferences and about preferences

- drop the separation between building objectives and optimization under preferences.
- reason with incomplete preferences (partial orders instead of total orders).

Search with preferences, but also through a space of preferences

- move to a completely other part of a large solution space by changing importance preferences (kind of Archimedean switch).

Emergent preference handling formalisms

These challenges and potential have led to different lines of work for handling preferences in AI, which had been carried out independently of each other. Examples are:

- **Non-monotonic logics with preferences**: this includes default logic allowing for reasoning with and about preferences as well as preferred answer-sets for extended logic programs.
- **Decision-theoretic planning**: this includes methods for plan evaluation based on expected utility as well as partially observable Markov decision processes.
- **Qualitative decision theory**: this includes graphical models for representing conditional ceteris paribus preference statements such as CP-networks.
- **Soft constraints**: this includes constraint evaluation frameworks based on semi-rings and work on constraint-based configuration with preferences.

At IJCAI 2001, we gained awareness that there may be an opportunity for an exchange.

Establishing preferences as an AI topic

Starting a new workshop

Why to invent the wheel again and again? Let us bring people together!

- We started with a **AAAI-02 workshop** on Preferences in AI and CP. I could win Francesca Rossi, Jim Delgrande, Jon Doyle, and Torsten Schaub as co-organizers.
- We published a **special issue of Computational Intelligence** on preferences in 2004 based on the success of this workshop.
- Gianni Bosi, Ronen Brafman, Jan Chomicki, and Werner Kiessling organized a **Dagstuhl-Seminar** on "Preferences: Specification, Inference, Applications" in 2004. It was a true multidisciplinary event gathering researchers from databases, AI, mathematics, decision science, philosophy who are all interested in the topic of preferences.

Based on the success of this seminar, the participants decided to start a **multidisciplinary workshop series**. Ronen and I were tasked to organize the first one at IJCAI-05.

M-PREF design choices

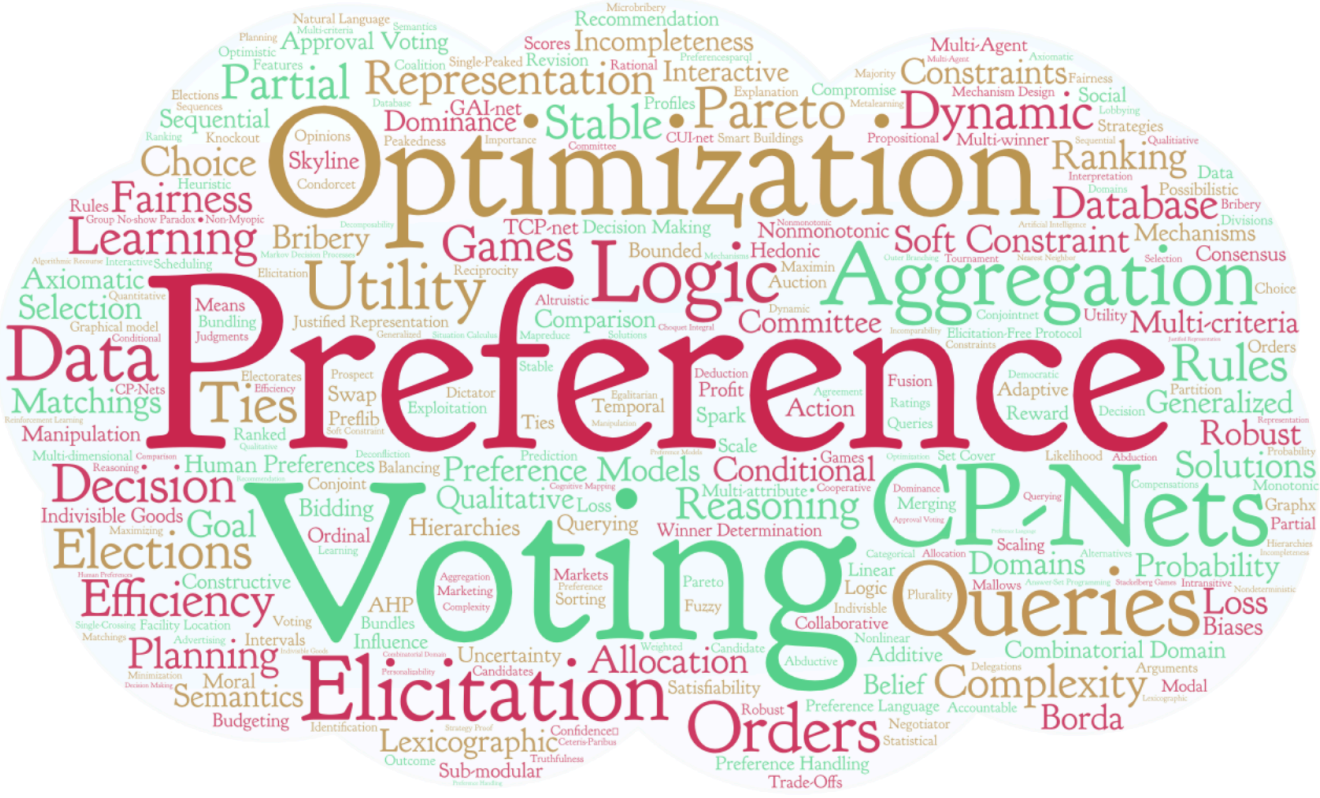
- **Multidisciplinary event**: facilitate exchange between researchers from different fields in AI, OR, and CS who are interested in preferences. This includes computational social choice, constraint satisfaction, data-bases, decision-theoretic planning, game theory, knowledge representation & reasoning, logic programming, machine learning, multi-agent systems, multi-criteria decision making.
- **Broad scope**: large range of topics, but with a focus on computational methods based on clear principles (avoid ad-hoc methods and pure mathematical models).
- **Multiplicity of tasks**: the term “handling” covers all kinds of computational tasks, namely elicitation, learning, modeling, representation, aggregation, and management of preferences as well as methods for reasoning about preferences.
- **Diverse environment**: changing geography and changing host conference. The last point was, however, abandoned as most participants came from the AI community.

New problems, new solutions!

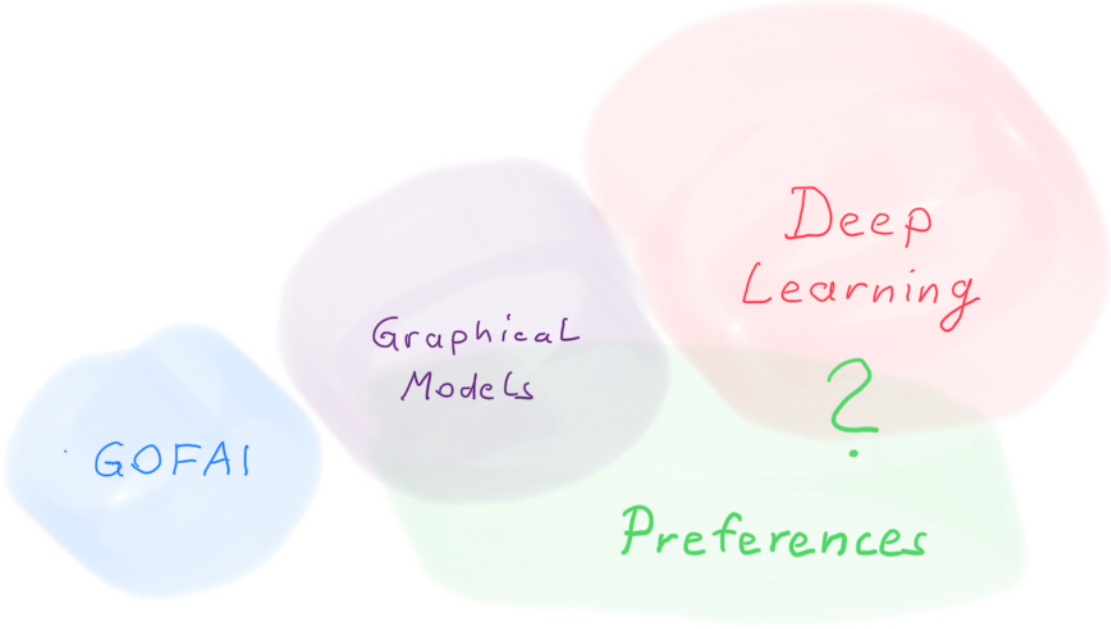
The M-PREF workshop series provides a forum to discuss new kinds of approaches and to increase their visibility:

- optimize under a set of utility functions in incremental preference elicitation.
- multiple preference relations in multi-agent settings and social choice.
- reasoning under qualitative preferences.
- preferential dependence of criteria.
- importance preferences between criteria.
- logical combinations of preference statements: negation, disjunction, quantification.
- new tasks: learning, explanation, relaxation, and revision of preferences. Reasoning with and about preferences.
- unbounded number of criteria (e.g. on components of complex systems).
- allow overriding of ceteris paribus preferences by more specific preferences.

Which topics have been dominant?



Will preferences remain important in AI?



Is the DM framework still valid in the age of ML?

Decision making: explore alternatives, compare them with each other, and choose a best one.

Machine learning: learn models that shortcut this process based on decisions from the past. However, these shortcuts may easily be invalidated by changes in the world:

- Changing alternatives
- Changing criteria
- Changing preferences

Those changes are shaping the decisions of the future. They will require a complete retraining of the ML model, which is costly and may cause more changes in the model than necessary.

The DM framework can model these changes in a direct way and describe which decisions are made under these changes. It thus respects the principle that *interesting changes in behavior must be expressible in a simple way* in (McCarthy, 1958).

Which topics might deserve more attention?

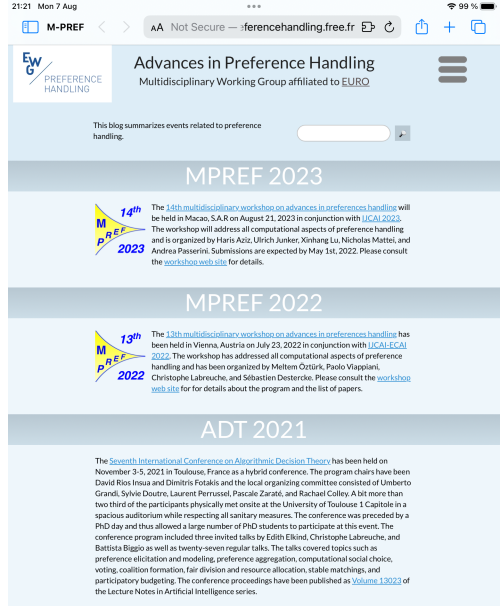
Here are my personal favorites:

- **Explanations and preferences**: there is more and more work on explaining learned preferences, but not so much on explaining decisions in terms of preferences.
- **Preference projection** from the criteria space to the decision space to help the optimization algorithm. Isn't backprop an example for this?
- **Preference-based machine learning**: can preferences help with multiple conflicting learning objectives and partially known objectives?
- **Preference change**: preferences are subject of choice as well and these choices may change as well. So there are no universal preferences, but there may be universal meta-preferences, which govern the revision of preferences.
- **Ethics and preferences**: the improvements that preference-based systems have for human users should not be paid with detriments for human trainers of those systems.

Future events on Advances in Preference Handling

subscribe to newsletter

visit web site



A few references

Discovery of preferences for AI: References

Problem solving vs. decision making

Simon, H. A. et al. Decision Making and Problem Solving. Interfaces 17, 11-31 (1987).

Taxonomic reasoning

Ginsberg, M. L. Readings in Nonmonotonic Reasoning. (Morgan Kaufman, 1987).

Temporal reasoning

Kautz, H. A. The logic of persistence. in AAAI Conference on Artificial Intelligence vol. 5 401–405 (1986).

Fault diagnosis

Junker, U. Preferring Diagnoses Using a Partial Order on Assumptions. Annals of Mathematics and Artificial Intelligence 11, 169–185 (1994).

Legal reasoning

Gordon, T. F. The Pleadings Game, An Artificial Intelligence Model of Procedural Justice. (Springer, 1995).

Prioritized defaults

Brewka, G. Preferred Subtheories: An Extended Logical Framework for Default Reasoning. in IJCAI'89: Proceedings of the 11th International Joint Conference on Artificial intelligence 1043–1048 (Morgan Kaufman, 1989).

Constraint optimization

Rossi, F., Beek, P. van & Walsh, T. Handbook of Constraint Programming. (Elsevier, 2006).

Challenges for preferences in AI: References

General framework for decision making

Keeney, R. L. & Raiffa, H. Decisions With Multiple Objectives: Preferences and Value Tradeoffs. (John Wiley and Sons, Inc., 1976).

Bouyssou, D., Marchant, T., Pirlot, M., Tsoukiàs, A. & Vincke, P. Evaluation and Decision Models with Multiple Criteria: Stepping stones for the analyst. (Springer, 2006).

Non-monotonic logics with preferences

Junker, U. A Cumulative-Model Semantics for Dynamic Preferences on Assumptions. in Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence vol. 1 162–167 (Morgan Kaufman, 1997).

Delgrande, J. P. & Schaub, T. Compiling Reasoning with and about Preferences into Default Logic. in Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence 168–174 (1997).

Brewka, G. & Eiter, T. Preferred answer sets for extended logic programs. Artificial Intelligence 109, 297–356 (1999).

Decision-theoretic planning

Blythe, J. Decision-Theoretic Planning. AI Magazine 20, 37–54 (1999).

Qualitative decision theory

Bacchus, F. & Grove, A. Graphical models for preference and utility. in Eleventh Conference on Uncertainty in Artificial Intelligence 3–10 (1995).

Doyle, J. & Thomason, R. H. Background to Qualitative Decision Theory. AI Magazine 20, 55–68 (1999).

Boutilier, C., Brafman, R. I., Domshlak, C., Hoos, H. H. & Poole, D. CP-nets: A Tool for Representing and Reasoning with Conditional Ceteris Paribus Preference Statements. Journal of Artificial Intelligence Research 21, 135–191 (2004).

Soft constraints

Bistarelli, S. et al. Semiring-Based CSPs and Valued CSPs: Frameworks, Properties, and Comparison. Constraints 4, 199–240 (1999).

Junker, U. & Mailharro, D. Preference programming: Advanced problem solving for configuration. Artificial Intelligence for Engineering Design, Analysis and Manufacturing 17, 13–29 (2003).

Establishing preferences as an AI topic: References

Starting a new workshop

Junker, U. Preferences in AI and CP: Symbolic Approaches. Papers from the AAAI Workshop. Technical Report WS-02-13. (AAAI Press, 2002).

Junker, U., Delgrande, J., Doyle, J., Rossi, F., Schaub, T., Preface[1]. Computational Intelligence 20, 109–110 (2004).

Bosi, G., Brafman, R. I., Chomicki, J. & Kießling, W. 04271 Abstracts Collection – Preferences: Specification, Inference, Applications. in Dagstuhl Seminar Proceedings (eds. Bosi, G., Brafman, R. I., Chomicki, J. & Kießling, W.) vol. 04271 1–17 (2006).

[1] to the special issue on preferences

New problems, new solutions!

Walsh, T. Representing and Reasoning with Preferences. AI Magazine 28, 59–70 (2007).

Doyle, J., Goldsmith, J., Junker, U. & Lang, J. Preference Handling for Artificial Intelligence. Papers from the AAAI Workshop. Technical Report WS-07-10. (AAAI Press, 2017).

Goldsmith, J. & Junker, U. Preference Handling for Artificial Intelligence. AI Magazine 29, 9 (2008).

Chomicki, J., Conitzer, V., Junker, U. & Perny, P. Advances in Preference Handling. Papers from the AAAI Workshop. Technical Report WS-08-09. (AAAI Press, 2008).

Brafman, R. I. & Domshlak, C. Preference Handling: An Introductory Tutorial. AI Magazine 30, 1–38 (2009).

Conitzer, V. Making Decisions Based on the Preferences of Multiple Agents. Communications of the ACM 53, 84–94 (2010).

Domshlak, C., Hüllermeier, E., Kaci, S. & Prade, H. Preferences in AI: An overview. Artificial Intelligence 175, 1037–1052 (2011).

Rossi, F., Venable, K. B. & Walsh, T. A Short Introduction to Preferences: Between AI and Social Choice. (Springer Cham, 2011).

Fürnkranz, J., Hüllermeier, E., Rudin, C., Slowinski, R. & Sanner, S. Preference Learning (Dagstuhl Seminar 14101). Dagstuhl Reports 4, 1–27 (2014).

Pigozzi, G., Tsoukiàs, A. & Viappiani, P. Preferences in artificial intelligence. Annals of Mathematics and Artificial Intelligence 77, 361–401 (2015).

Will preferences remain important in AI: References

McCarthy, J. Programs with common sense. in Teddington Conference on the Mechanization of Thought Processes (1958).

Explanations and preferences

Junker, U. Preference-based Problem Solving for Constraint Programming. in CSCLP 2007 (eds. Fages, F., Rossi, F. & Rossi, F.) 109–126 (Springer-Verlag, 2008).

Belahcene, K., Labreuche, C., Maudet, N., Mousseau, V. & Ouerdane, W. Explaining robust additive utility models by sequences of preference swaps. Theory and Decision 82, 151–183 (2017).

Preference Projection

Junker, U. Preference-Based Search and Multi-Criteria Optimization. Annals of Operations Research 130, 75–115 (2004).

Preference-based machine learning

Désidéri, J.-A. Multiple-gradient Descent Algorithm for Pareto-Front Identification. in Modeling, Simulation and Optimization for Science and Technology (eds. Fitzgibbon, W., Kuznetsov, Y. A., Neittaanmäki, P. & Pironneau, O.) vol. 34 41–58 (Springer, 2014).

Preference change

Grüne-Yanoff, T. & Hansson, S. O. Preference Change: Approaches from Philosophy, Economics and Psychology. (Springer, 2009).

Junker, U. Upside-down Preference Reversal: How to override ceteris-paribus preferences? in vol. 9060 186–201 (Springer International Publishing, 2015).

Ethics and preferences

Rowe, N. ‘It’s destroyed me completely’: Kenyan moderators decry toll of training of AI models (The Guardian, Wed 2 Aug 2023).