1.3 Typical examples

1.3.1 The excursion problem after Chernoff (1959)

- Mr. Nelson wants to go on a mountain hike tomorrow, at a at a time of year when the sudden arrival of bad weather is to be expected.

- **action options:**

$$\begin{aligned} a_1 &:= \text{carry light clothing} \\ a_2 &:= \text{take light clothing plus umbrella} \\ a_3 &:= \text{take weatherproof, warm clothing plus umbrella} \end{aligned}$$

- The **relevant facts** or **conditions** are.

$$\begin{aligned} \vartheta_1 &:= \text{nice weather on the day of the trip} \\ \vartheta_2 &:= \text{bad weather on the day of the trip} \end{aligned}$$

- utility table for Mr. Nelson:

	ϑ_1	ϑ_2
a_1	5	0
a_2	3	1
a_3	2	3

Valuations are interpreted as utility.

Potential data-based additional info ($\rightarrow$ later): consult a barometer

1.3.2 Participation in a lottery

Given an urn with

g	green
b	blue
r	remaining

Balls. One can

a_1	not play
a_2	bet on green at the price of c_g
a_3	bet on blue at the price of c_b.

Now a ball is drawn at random. If it is green (resp. blue), you get w_g and w_b (w: win), respectively, if you betted on green (resp. blue).

Let the amounts be small enough to be able to assume linearity of utility.

$\theta = \{\vartheta_1, \vartheta_2, \vartheta_3\}$ with

ϑ_1 drawn ball is green
ϑ_2 drawn ball is blue
ϑ_3 drawn ball is other color

	ϑ_1	ϑ_2	ϑ_3
a_1	0	0	0
a_2	$w_g - c_g$	$-c_g$	$-c_g$
a_3	$-c_b$	$w_b - c_b$	$-c_b$

Important for later: If g,b,r are known, we are in an uncertainty situation of Type I (risk situation)!

1.3.3 „Dividing a cake“

Given is a cake consisting of 8 numbered pieces of the same size.

- The decision maker chooses an edge between two pieces z and $z+1$, $z \in \{1, \ldots 7\}$, where the cake is shared.

- There are two states of nature: ϑ_1 and ϑ_2. If ϑ_1 occurs then the decision maker receives the pieces 1 to z, while if ϑ_2 occurs (s)he gets the pieces $z+1$ to 8.
utility $=$ number of pieces! (Non-saturation assumption)

	ϑ_1	ϑ_2
1	1	7
2	2	6
3	3	5
4	4	4
5	5	3
6	6	2
7	7	1

- Nothing is presupposed here yet, how the states of nature occur. If they are generated e.g. by an adversary, then we get a typical example for the situation of the strategic game, i.e. for Type II-uncertainty.

1.3.4 Investment problem

A company is faced with the decision to make a marketing investment. Its success depends on the state of the economy in the next six months.

- **actions**

$a_1 :=$ make investment.

$a_2 :=$ do not make investment

- **states**

$\vartheta_1 :=$ rising economy

$\vartheta_2 :=$ stagnation

$\vartheta_3 :=$ economic activity falling

	ϑ_1	ϑ_2	ϑ_3
a_1	10000	2000	-15000
a_2	1000	1000	0

1.3.5 Production planning under scenarios

- q goods
- $\mathbb{A} \subseteq (\mathbb{R}_0^+)^q$ described by capacity constraints.
- $$a = \begin{pmatrix} a[1] \\ \vdots \\ a[q] \end{pmatrix}$$

  with $a[\ell]$: Production of $a[\ell]$ units of good ℓ, $\ell = 1, \dots, q$
- state of nature ϑ: scenario (possible development of the market)

- sales under different scenarios result in utility function:

$$\begin{aligned} u : \mathbb{A} \times \Theta &\to \mathbb{R} \\ (a, \vartheta) &\mapsto u(a, \vartheta) \end{aligned}$$

- Simple form: unit price based, under non-saturation assumption: vector $c(\vartheta)$ with $c[\ell](\vartheta)$: Unit price for good ℓ under scenario ϑ, $\ell = 1, \ldots, q$, then

$$u(a, \vartheta) = a'c(\vartheta) = \sum_{\ell=1}^{q} a[\ell] \cdot c[\ell](\vartheta)$$

or, if $c(\vartheta)$ is interpreted as a cost:

$$\mathsf{l}(a, \vartheta) = \sum_{\ell=1}^{q} a[\ell] \cdot c[\ell](\vartheta)$$

1.3.6 Stock purchase (highly simplified, of course)

1.3.7 Embedding statistical hypotheses tests into decision theory I: the data-free core of a testing problem

1.3.8 Embedding parameter estimation into decision theory I: the data-free core of an estimation problem

1.3.9 Embedding supervised learning: the data-free core of a classification task

1.3.10 The liberal view – preliminary remarks

By „daring“ an often purely technical reinterpretation of the states of nature, a number of other problems can be analysed in terms of decision theory.

However, it is crucial to check the conditions on the semantics of a decision problem from Remark 1.6 before drawing conclusions from this analogy/ embedding.

Typical examples could include

- dominance considerations $\longrightarrow$ admissibility, Pareto front.
- Worst-case scenario analysis (cf. maximin criterion).
- Weighting of states of nature (cf. Bayes criterion).

1.3.11 The liberal view I: the decision-theoretic account of algorithm selection via benchmark experiments

The two graphs are taken from:
Antonucci, A. & Corani, G. (2017):The multilabel naive credal classifier. *International Journal of Approximate Reasoning* **83**: 320-336, `https://doi.org/10.1016/j.ijar.2016.10.006`

Algorithms:

- multilabel naive credal classifier (MNCC).
- naive credal classifier (NCC)
- multilabel naive Bayes classifier (MNBC)
- naive Bayes classifier (NBC)

Table 4
Marginal performance with ranks (in the parentheses) of the algorithms.

#	Data Set	u_{80}		Accuracy	
		MNCC (2)	NCC (1.8)	MNBC (3.3)	NBC (2.8)
1	Emotions	77.82 (2)	78.03 (1)	76.95 (4)	77.09 (3)
2	Scene	83.86 (1)	82.92 (3)	83.35 (2)	82.74 (4)
3	Flags	77.98 (1)	77.32 (2)	74.06 (4)	74.29 (3)
4	E-mobility	87.03 (3)	88.42 (1)	86.82 (4)	88.32 (2)
5	Yelp	82.22 (2)	82.28 (1)	79.57 (4)	80.04 (3)
6	Birds	87.77 (2)	87.91 (1)	84.33 (4)	84.69 (3)
7	Genbase	99.29 (3)	99.28 (4)	99.90 (1)	99.89 (2)
8	Yeast	69.01 (2)	69.87 (1)	67.84 (4)	68.89 (3)
9	Medical	93.54 (3)	93.51 (4)	94.16 (2)	94.62 (1)
10	Slashdot	94.50 (2)	94.64 (1)	93.90 (4)	94.06 (3)
11	Nus-Wide	78.10 (1)	77.79 (2)	77.05 (3)	76.87 (4)
12	Imdb	88.76 (2)	89.18 (1)	87.63 (4)	88.47 (3)
13	Enron	81.06 (2)	81.09 (1)	77.69 (4)	77.84 (3)
14	Mediamill	78.10 (1)	77.79 (2)	77.05 (3)	76.87 (4)
15	Cal500	72.87 (1)	72.25 (2)	70.03 (4)	70.08 (3)

Source: Antonucci & Corani (2017, IJAR, p. 330).

A. Antonucci, G. Corani / International Journal of Approximate Reasoning 83 (2017) 320–336

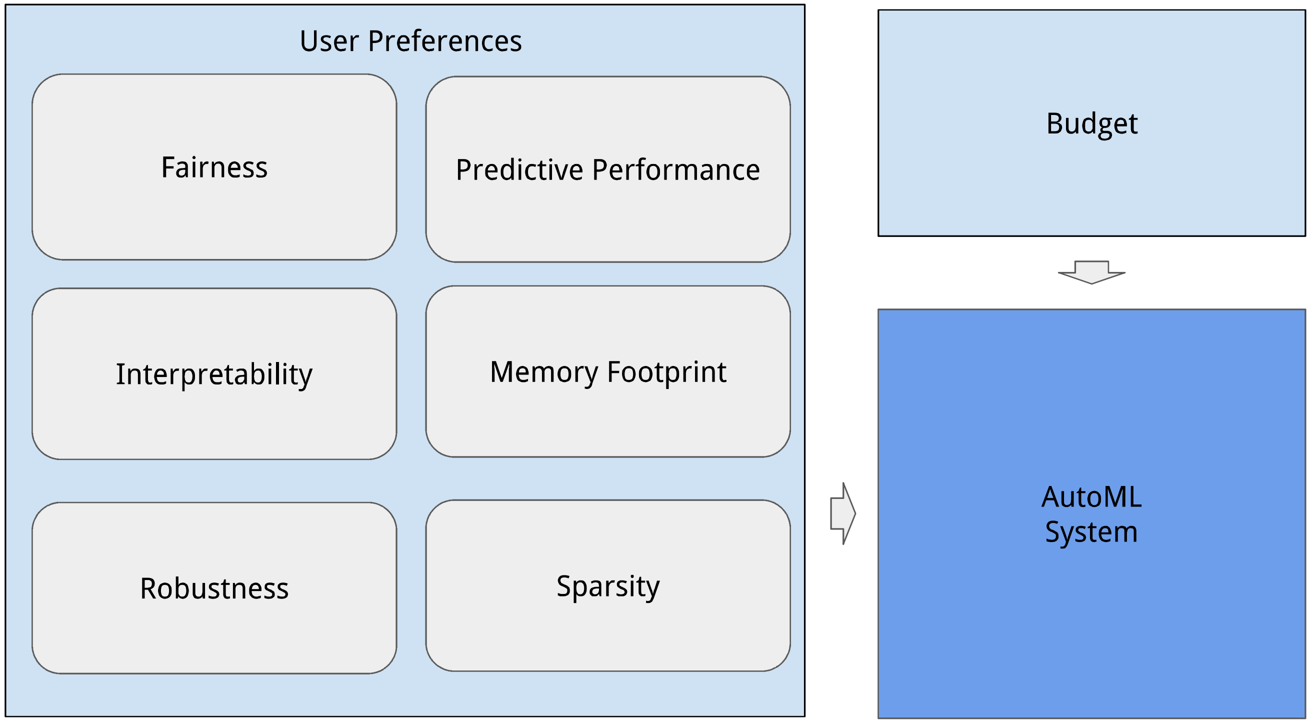
Table 2
Characteristics of the data sets.

#	Data set	Ref.	Classes	Density	Features	Instances
1	Emotions	[43]	6	0.311	72/44	593
2	Scene	[8]	6	0.179	294/224	2407
3	Flags	[30]	7	0.484	19/14	194
4	E-mobility	[10]	8	0.121	13/12	3218
5	Yelp	[39]	8	0.295	668/203	1951
6	Birds	[9]	11	0.116	260/90	435
7	Genbase	[26]	13	0.087	1185/68	662
8	Yeast	[27]	13	0.325	103/92	2417
9	Medical	[36]	14	0.084	1449/292	888
10	Slashdot	[38]	14	0.084	1079/462	3663
11	Nus-Wide	[12]	16	0.114	498/408	1981
12	Imdb		19	0.097	1001/590	8792
13	Enron	[33]	24	0.131	257/214	1696
14	Mediamill	[41]	25	0.157	120/91	4857
15	Cal500	[44]	119	0.201	68/67	502

Source: Antonucci & Corani (2017, IJAR, p. 329).

1.3.12 The liberal view II: multi-criteria problems and the decision-theoretic account of some Auto ML issues

The following graph is taken from:
Pfisterer, F., Coors, S., Thomas, J. & Bischl, B. (2019/21): Multi-Objective Automatic Machine Learning with AutoxgboostMC. *Automating Data Science Workshop at ECML 2019*, `https://arxiv.org/abs/1908.10796` [v2 (2021), last access April 23rd, 2023]



Source: Pfisterer, Coors, Thomas & Bischl (2019, arxiv, p. 4).

1.4 Randomized actions

.

Def. 1.7. [Randomized (mixed actions)]

Let $\mathbb{A}$ be the action set of a data-free decision problem $(\mathbb{A}, \Theta, u(\cdot))$ (and $\sigma(\mathbb{A})$ a σ-algebra over $\mathbb{A}$ that contains all one-point sets $\{a\} \in \mathbb{A}$.)
Then every probability measure on $(\mathbb{A}, \sigma(\mathbb{A}))$ is called randomized (mixed) action.
The set of all randomized actions on $(\mathbb{A}, \sigma(\mathbb{A}))$ is denoted by $\mathcal{M}^{\sigma(\mathbb{A})}(\mathbb{A})$. When it is evident which σ-algebra is used, one writes briefly $\mathcal{M}(\mathbb{A})$.

Rem. 1.8. [On the semantics of a randomized action]

Let $\mathbb{A} = \{a_1, \ldots, a_n\}$. Then the randomized action $\widetilde{a}(\cdot) \in \mathcal{M}(\mathbb{A})$ consists of the following action rule:
Run a random experiment on $\{1, \ldots, n\}$ with $p(\{i\}) = \widetilde{a}(\{a_i\}), \quad i = 1, \ldots, n$, and choose action a_i exactly when i occurs.
That is, with probability $\widetilde{a}(\{a_i\})$ the action a_i is chosen.
One often writes

$$\widetilde{a} = \begin{bmatrix} a_1 & \ldots & a_n \\ \widetilde{a}(\{a_1\}) & \ldots & \widetilde{a}(\{a_n\}) \end{bmatrix}.$$

If one chooses $\mathcal{M}(\mathbb{A})$ as the action set, then one can formulate a corresponding decision problem. For this it is still necessary to define the utility / loss appropriately (naturally as an expected value).

Def. 1.9. [Mixed extension]

Given a data-free decision problem $(\mathbb{A}, \Theta, u(\cdot))$ and a suitable σ-algebra $\sigma(\mathbb{A})$ on $\mathbb{A}$. Then the data-free decision problem $(\mathcal{M}^{\sigma(\mathbb{A})}(\mathbb{A}), \Theta, \widetilde{u}(\cdot))$ with

$$
\begin{aligned}
\widetilde{u}(\cdot): \quad \mathcal{M}^{\sigma(\mathbb{A})} \times \Theta &\longrightarrow \mathbb{R} \\
(\widetilde{a}, \vartheta) &\longmapsto \widetilde{u}(\widetilde{a}, \vartheta)
\end{aligned}
$$

and

$$
\widetilde{u}(\widetilde{a}, \vartheta) := \mathbb{E}_{\widetilde{a}}[u(a, \vartheta)] = \int u(a, \vartheta) \; \mathrm{d}\widetilde{a}(a) \tag{1.6}
$$

is called the <u>mixed extension</u> of $(\mathbb{A}, \theta, u(\cdot))$ (with respect to $\sigma(\mathbb{A})$).

Rem. 1.10. [On Def. 1.9]

- The use of the general measure integral allows the simultaneous consideration of the continuous and discrete case as well as the consideration of mixed continuous/discrete distributions. In particular:
  If $\widetilde{a}(\cdot)$ is a (Lebesgue) continuous probability measure with density $\widetilde{f}(\cdot)$, then

$$\widetilde{u}(\widetilde{a}, \vartheta) = \int_{\mathbb{A}} u(a, \vartheta)\widetilde{f}(a)\, da\,. \tag{1.7}$$

  If $\widetilde{a}(\cdot)$ is discrete with probability function $\widetilde{p}(\cdot)$ on $\mathbb{A} = \{a_1, a_2, \ldots, \}$, then

$$\widetilde{u}(\widetilde{a}, \vartheta) = \sum_{i=1}^{n} u(a_i, \vartheta)\widetilde{p}(a_i)\,. \tag{1.8}$$

- Note formula (1.6) is a definition. It is plausible, but by no means logically compelling.

- Thus, the mixed extension is again a data-free decision problem. Formulating definitions and propositions, there is typically no need to distinguish whether we look at an „original decision problem“ or the corresponding mixed extension.

Rem. 1.11. [Pure actions]

The randomized actions of the form $\widetilde{a}(\cdot) = \delta_{\{a\}}, a \in \mathbb{A}$ (Dirac measure = single point measure in the set $\{a\}$) are called pure actions and identified with $a \in \mathbb{A}$.

Rem. 1.12. [On the sense and nonsense of randomized actions]