



THE UNIVERSITY
of ADELAIDE

Markus Wagner

Kinder Surprise's Debut in Discrete Optimisation – A Real-World Toy Problem that can be Subadditive

<https://github.com/markuswagnergithub/TOYlib>
markus.wagner@adelaide.edu.au





THE UNIVERSITY
of ADELAIDE

Markus Wagner

Kinder Surprise's Debut in Discrete Optimisation - A Real-World Toy Problem that can be Subadditive

<https://github.com/markuswagnergithub/TOYlib>
markus.wagner@adelaide.edu.au





(Creative Commons Photo)

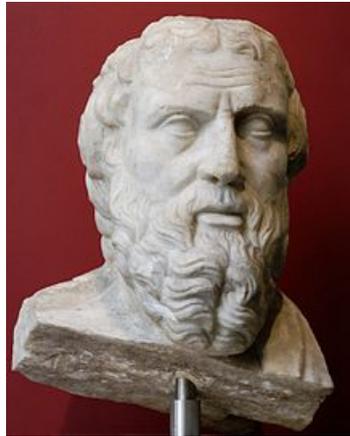
Some time in 2018, near Berlin...

From a discussion:

- “submodular functions”
- “subadditive/superadditive functions”
- “curvature”
- “diminishing returns”
- “matroid constraints”
- “approximation ratios”
- “social welfare”
- “real-world examples”



"The whole is more than the sum of its parts"



"The whole is more than the sum of its parts"



...Thanks, Aristotle, you've got any data to support this?

Some time in 2018, near Berlin...

- Kinder Surprise: shops sell the complete sets for less than the sum
- Stamp Collections: no price lists online
- "profit in manufacturing": iPhone costs \$500 to produce, then add marketing, development, ... and somewhere between \$1500 and \$500 Apple is still making money → "many parts to a few products" relationship, and data is super scarce and incomplete.
- LEGO sets "many components to many products"
Lists of parts and prices of parts online, automation is not trivial, and one cannot always find perfect matches
- algorithm portfolios: some data online, but only up to $n=20$ or so solvers
- specialisation of other kinds: think of human moving from "general purpose hunters and gatherers" to "bakers, Beyoncé, and the BB-DOB organisers". Data? Nope.
- "The whole is more than the sum of its parts" – Thanks, Aristotle, you've got any data to support this?

TOYlib

github.com/markuswagnergithub/

TOYlib

Each figurine:

- Anonymised name
- Price
- Set ID
- Low-resolution photo
- Relatable

187 sets of figurines produced between 1979 and 2018: 2366 items in total



Andre Feiler from the Feiler Verlag feiler-verlag.de, Germany kindly provided us with a digital copy of the current pricing guide

Creation of the Dat

11 General Hati 9,00 €

12 Kleiner Elefant (Schwanz oft abgebrochen!) 10,00 €

Beipackzettel deutsch 170,00 €

Beipackzettel englisch (Jungle-Book) 150,00 €

Title

Dschungelbuch D 1985 (Varianten) (© Walt Disney Productions)

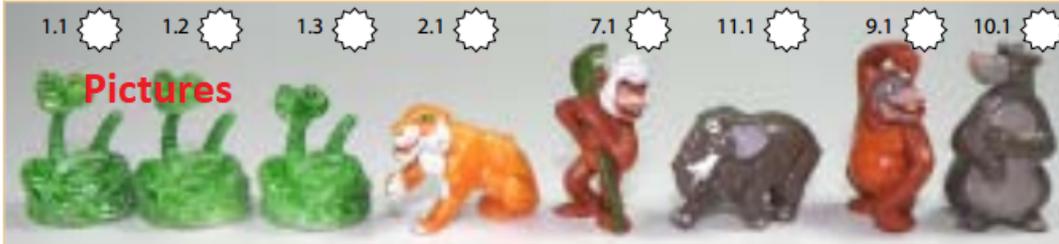


		Table		
1.1	Kaa (sattgrüne Bemalung)			45,00 €
1.2	Kaa (hellgrüne Bemalung)			60,00 €
1.3	Kaa (kurzer Hals, resultiert aus Gussfehler bei der Produktion)	7.2	7.3	260,00 €
2.1	Shir Khan (hellorange Bemalung)			15,00 €
7.1	Diener (Zähne weiß bemalt)			15,00 €
7.2-7.3	Diener (rechter/linker Fuß mit weit übersteh. Gussabdruck)			200,00 €
9.1	King Louis (Zähne weiß bemalt)			15,00 €
10.1	Balu (Tatzen bemalt)			15,00 €
11.1	General Hati (Fußnägel bemalt)			15,00 €



Serienzubehör Dschungelbuch		
1	Malkasten (unbenutzt)	75,00 €
2	Quartett (Seiten lose)	50,00 €
2.1	Quartett (oben gebunden)	120,00 €
2.2	Quartett (seitl. gebunden)	140,00 €

1.1 Kaa (sattgrüne Bemalung) 45,00 €

1.2 Kaa (hellgrüne Bemalung) 60,00 €

1.3 Kaa (kurzer Hals, resultiert aus Gussfehler bei der Produktion) 260,00 €

2.1 Shir Khan (hellorange Bemalung) 15,00 €

7.1 Diener (Zähne weiß bemalt) 15,00 €

7.2-7.3 Diener (rechter/linker Fuß mit weit übersteh. Gussabdruck) 200,00 €

9.1 King Louis (Zähne weiß bemalt) 15,00 €

10.1 Balu (Tatzen bemalt) 15,00 €

11.1 General Hati (Fußnägel bemalt) 15,00 €

Dschungelbuch D 1985 (Varianten) (© Walt Disney Productions)

Dschungelbuch GB 1990 (© Walt Disney Productions)

Serienzubehör Dschungelbuch

1 Malkasten (unbenutzt) 75,00 €

2 Quartett (Seiten lose) 50,00 €

2.1 Quartett (oben gebunden) 120,00 €

Creation of the Data Set

Hartplastik - Serienfiguren 1981 bis heute

Olympiade der Schlümpfe D 1983 (Varianten Zubehör)

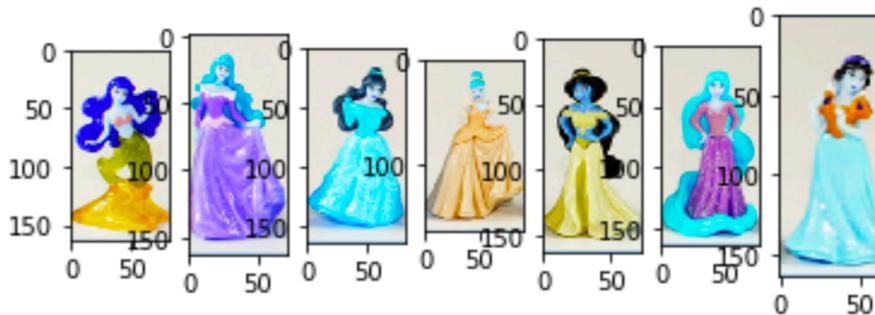
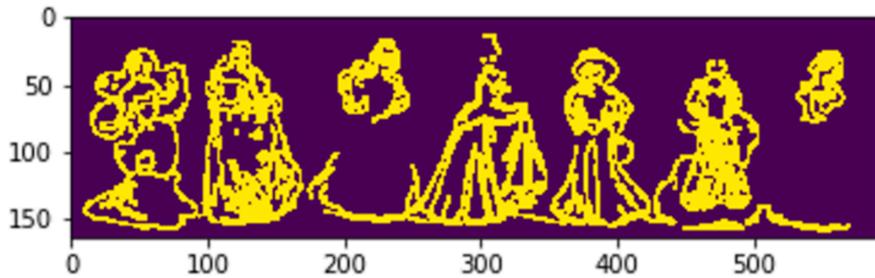
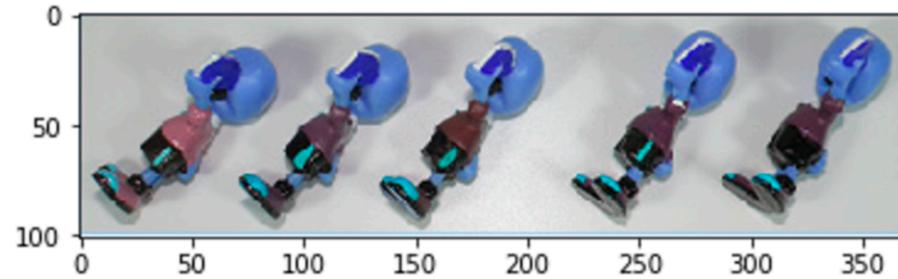


Farbunterschiede Zubehör „Olympiade der Schlümpfe“

A	Stelzen - links hellbraun, rechts dunkelbraun	
B	Hüpfseil - links dünnes Seil, rechts dickes Seil, mind. 3 Farbtöne gut unterscheidbar	
C	Löffel - links hellbraun, rechts dunkelbraun	
D	Ei - links hellrot, rechts dunkelrot	
7a	Stelzen, original am Steg	1500,00 €
1a	Löffel, mit Steg (incl. Ei)	1500,00 €



Creation of the Data Set



Steps in the image extraction process. From top to bottom: original image, non-tuned edge detection, final result after tuning. Units are pixels.

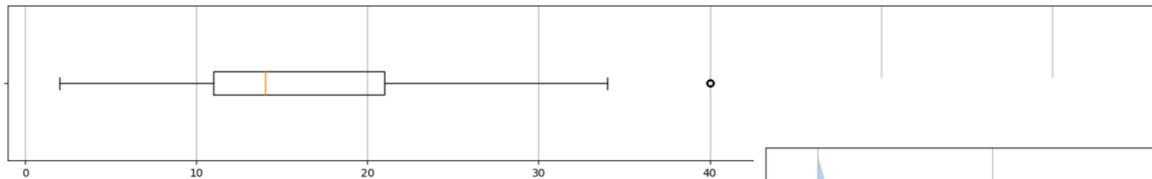
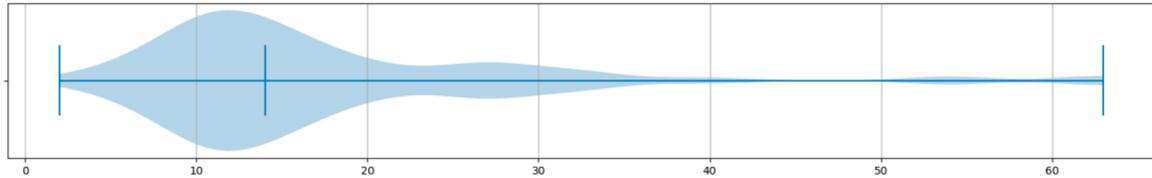
Creation of the Data Set

Excerpt from `KinderSurprise2019data.csv`. For example, the last set with the number 187 contains 10 figures (each valued at EUR 3.00) and 1 leaflet (valued at EUR 0.50).

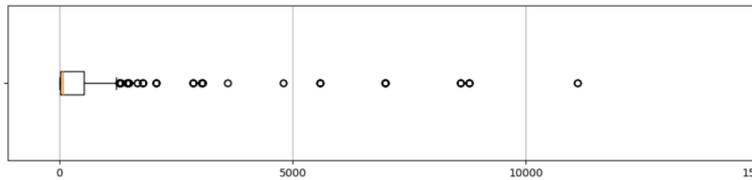
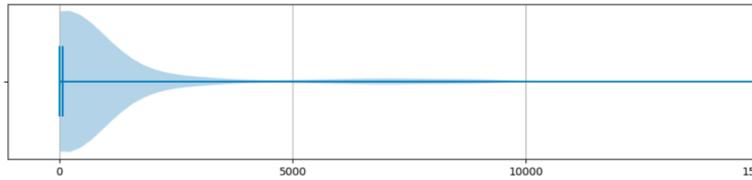
item_overall_id	item	item_in_set	item_price	item_image_prefix	set	set_size	set_total_price
1	1_1	1	1200	2_1	1	4	4800
2	1_2	2	1200	2_2	1	4	4800
3	1_3	3	1200	2_3	1	4	4800
4	1_4	4	1200	2_4	1	4	4800
5	2_1	1	200	2_5	2	3	600
6	2_2	2	200	2_6	2	3	600
7	2_3	3	200	2_7	2	3	600
				⋮			
2362	187_7	7	3	243_7	187	11	30.5
2363	187_8	8	3	243_8	187	11	30.5
2364	187_9	9	3	243_9	187	11	30.5
2365	187_10	10	3	243_10	187	11	30.5
2366	187_11	11	0.5	243_11	187	11	30.5

Data Characterisation 1/2

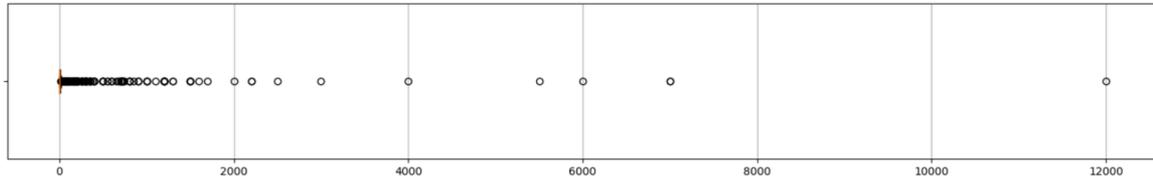
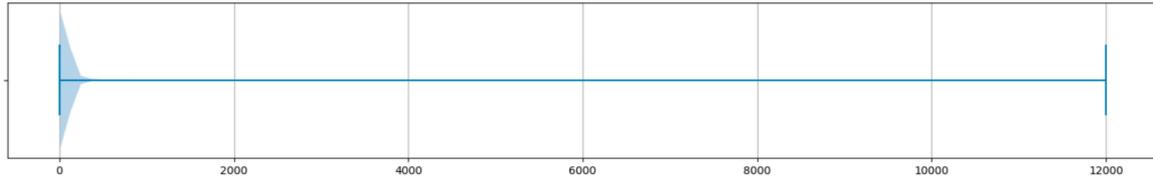
set_size



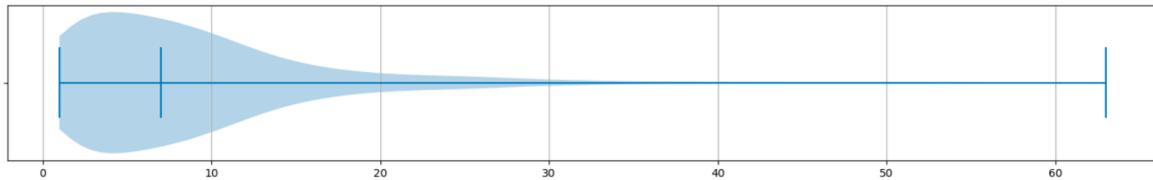
set_total_price



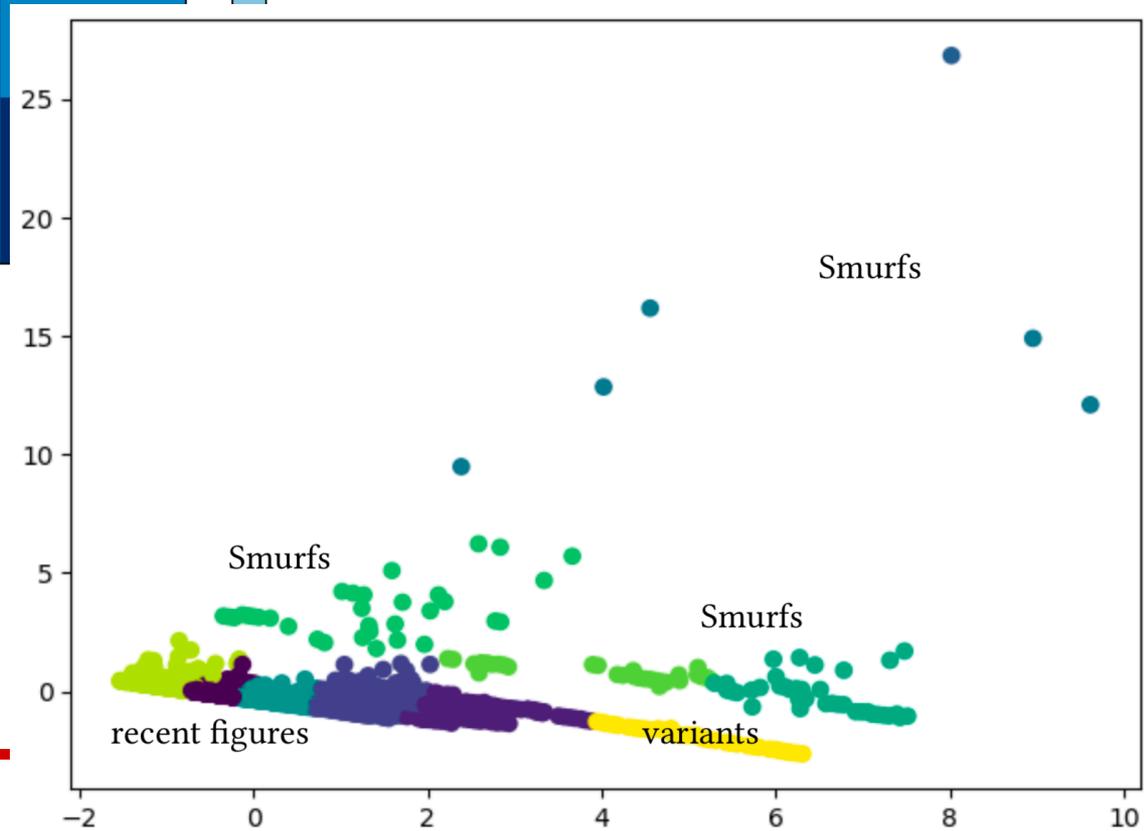
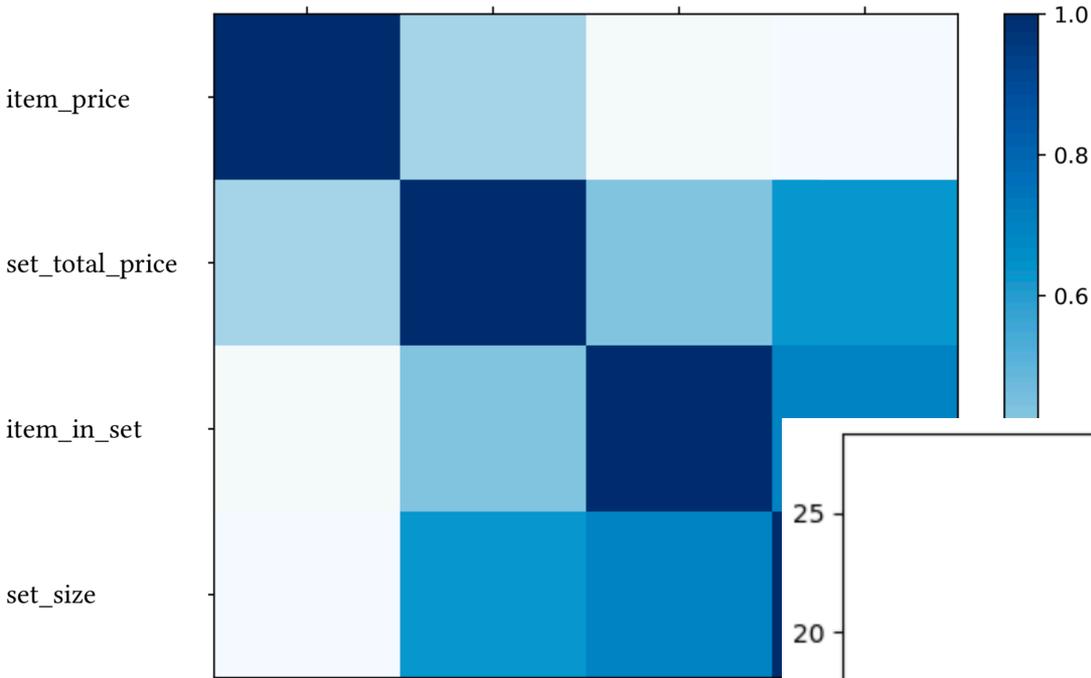
item_price



item_in_set



Data Characterisation 2/2





<https://static.adweek.com/adweek.com-prod/wp-content/uploads/2018/10/WhatsNextRoadSign.jpg>

Optimisation Problems 1/3

Functions with Matroid Constraints

Matroid constraints (if uniform: cardinality constraints) → When collecting figurines, we can easily think of scenarios where the task is to **distribute a fixed number of figurines among collectors** – hence the connection to functions with cardinality constraints.

Submodular functions: diminishing returns (strictly speaking: non-increasing) → Diminishing returns can occur when the happiness function of a collector contains aspects such as **“the increase of happiness is sub-linear with the number of figurines a collector has”**.

1977: greedy algorithm achieves a $1/2$ approximation ratio when maximizing monotone submodular functions under partition matroid constraints → **How close to algorithms get in the real world?** Data sets are surprisingly rare...

2019: **social welfare problem, n players compete for m items**, where items can have different values (“utilities”) for each player → Friedrich et al. then defined the constraint that each item can be allocated only to one player, and that a function f that, when maximizing f , is equivalent to maximizing a monotone function under a partition matroid constraint.

Optimisation Problems 2/3

Combinatorial Auctions

Single-minded bidders [23] – Here, the bidder is only happy if she gets exactly the items that she is interested in. If a bidder's wishes are not fulfilled (even just partially), then she is not happy. → (1) can be NP-hard, (2) for our case, having single-minded bidders makes the problem neither subadditive nor superadditive due to the happiness that is essentially a needle-in-a-haystack (also not submodular)

OR-defined happiness [23] – A bit more lenient... There, bids are fulfilled – and thus contribute to happiness – if at least one item from a set of desired items is assigned. In addition, a bidder can specify multiple such sets. The overall happiness here can then be defined as the sum of the individual happiness-es assigned to each fulfilled set.

[23] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. Algorithmic Game Theory, 2007.

Optimisation Problems 2/3

Combinatorial Auctions



New: (un-)capped happiness

More natural (?): **bidders get happier and happier, but also with diminishing returns.**

→ OR-bids + mathematical series.

Assuming n bidders, where each bidder i desires k items $d_{i,1}..d_{i,k}$ (k can vary across bidders):

- For an allocation of items to bidders, sort for each bidder individually the set of items that she actually gets according to their value, largest first, resulting in $g_{i,1}..g_{i,m}$. Undesired but assigned items are simply ignored.

- Optimisation Problem Variant 1 “uncapped happiness”:

$$\text{Happiness}(i) = g_{i,1}.\text{value} * 1 + g_{i,2}.\text{value} * 1/2 + g_{i,3}.\text{value} * 1/3 + \dots = \sum_{j=1}^k 1/j * g_{i,j}.\text{value}$$

- Optimisation Problem Variant 2 “capped happiness”:

$$\text{Happiness}(i) = g_{i,1}.\text{value} * 1 + g_{i,2}.\text{value} * 1/4 + g_{i,3}.\text{value} * 1/9 + \dots = \sum_{j=1}^k 1/j^2 * g_{i,j}.\text{value}$$

- The total sum across all bidders is then the total sum of each bidder’s happiness.

Variant 2: looks like BinVal (linear function with “binary values” as coefficients), where decision variables have the potential to dominate the additive effect of the other decision variables. Despite this, the variant remains a linear function.

Optimisation Problem 3/3

Uncharted Territory: Breaking submodularity and similar properties

Interesting direction for both theoretical and empirical studies:

- define functions that are almost everywhere submodular/subadditive/... **but with a few sets or points, where this property breaks**
- Give rise to interesting statements for mutation operators and greedy algorithms with restarts?
- To the best of our knowledge, no works about such occasional violations or prevailing conditions exist yet – and in particular about their effects – for example, about “almost always submodular functions”. While there are works characterising, e.g., the degree of submodularity via the so-called curvature [7, 30], they do not consider violations.

For this, data is again surprisingly scarce... or is there a parameterisable OneMax-equivalent out there?

Optimisation Problem 3/3

Unchartered Territory: Breaking submodularity and similar properties

Kinder Surprise (again) → (some) sets priced higher than the sum of the individual components? What intuitively might make sense at first sight backfires quickly for vendors, as buyers just have to purchase the individual items in order to get an eventual “boost” for free. In principle, this might still work for rare collections where hardly anyone has individual components, so that having the full set is extremely rare and nobody wants to part with the figurines for little money and sentimental reasons.

Profit in manufacturing → not so sure... profit?

LEGO sets → similar to Kinder Surprise: do we favour the collectors or the vendors?

Algorithm portfolios → cannot be super-additive

Aristotle? → no data on super-additivity

Summary



TOYlib:

- → relatable, discrete optimisation problems!
- the Kinder Surprise 2019 data set forms the beginning
- <https://github.com/markuswagnergithub/TOYlib> open source! GPL3
- acquire additional data sets with a particular focus on collectables that are organised in sets (trading cards from sports/card games, and more serious domains such as the collection of stamps)
- Interfaces to optimisation packages and frameworks like IOHprofiler?
- Also online: Jupyter notebook for Gurobi, i.e., a basic example for mathematical optimisation

We would like to thank Lujun Weng for his technical support, Tobias Friedrich for coining the term “real-world toy problem”, and we would like to thank Andreas Göbel, Timo Kötzing, and Francesco Quinzan for their discussions in the overall project.