

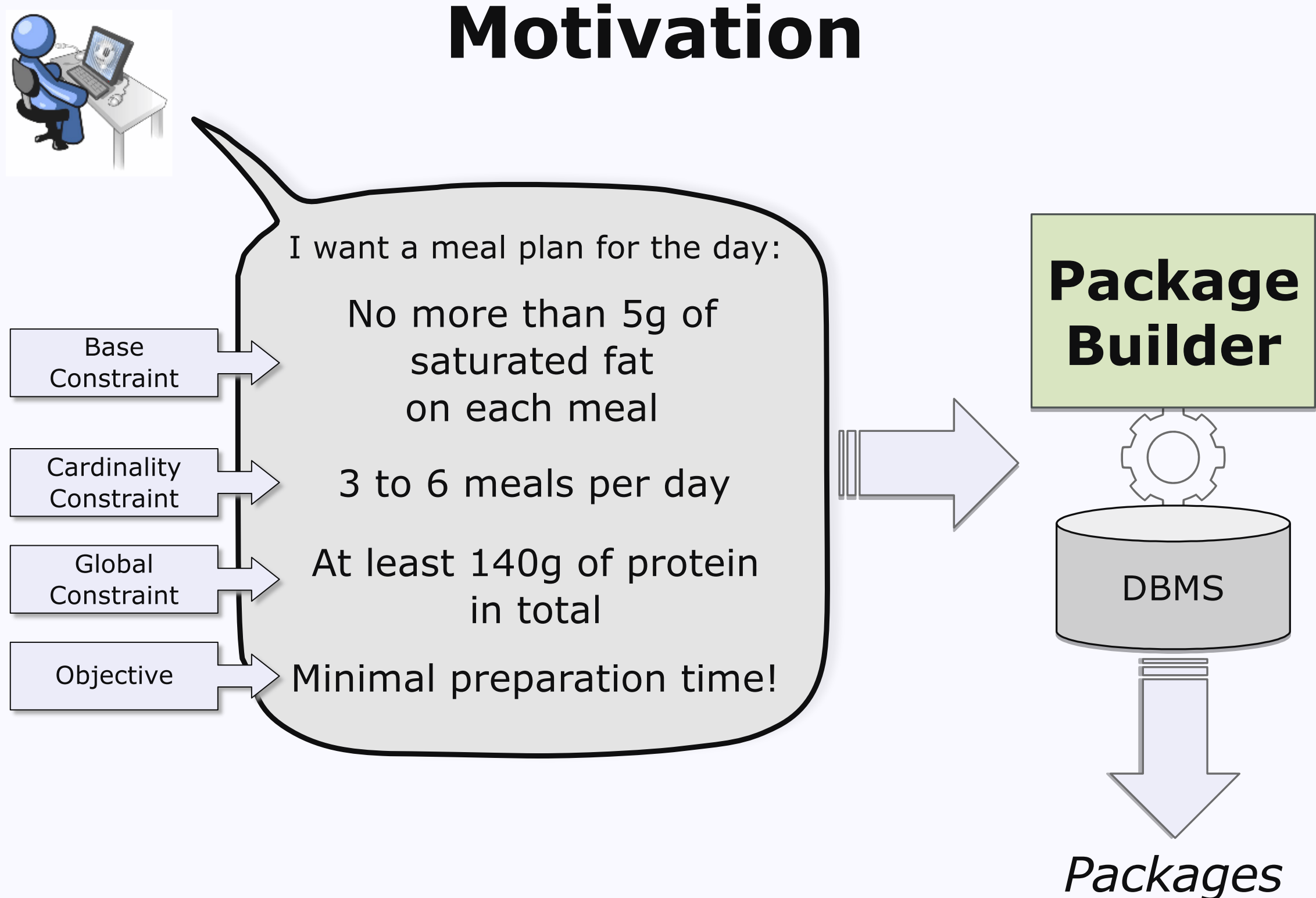
PackageBuilder: From Tuples to Packages

Matteo Brucato*, Rahul Ramakrishna*, Azza Abouzied[§] and Alexandra Meliou*
*UMass Amherst [§]New York University, Abu Dhabi

<http://packagebuilder.cs.umass.edu>



Motivation



PackageBuilder

PackageBuilder is a system that extends query engines to support package generation.

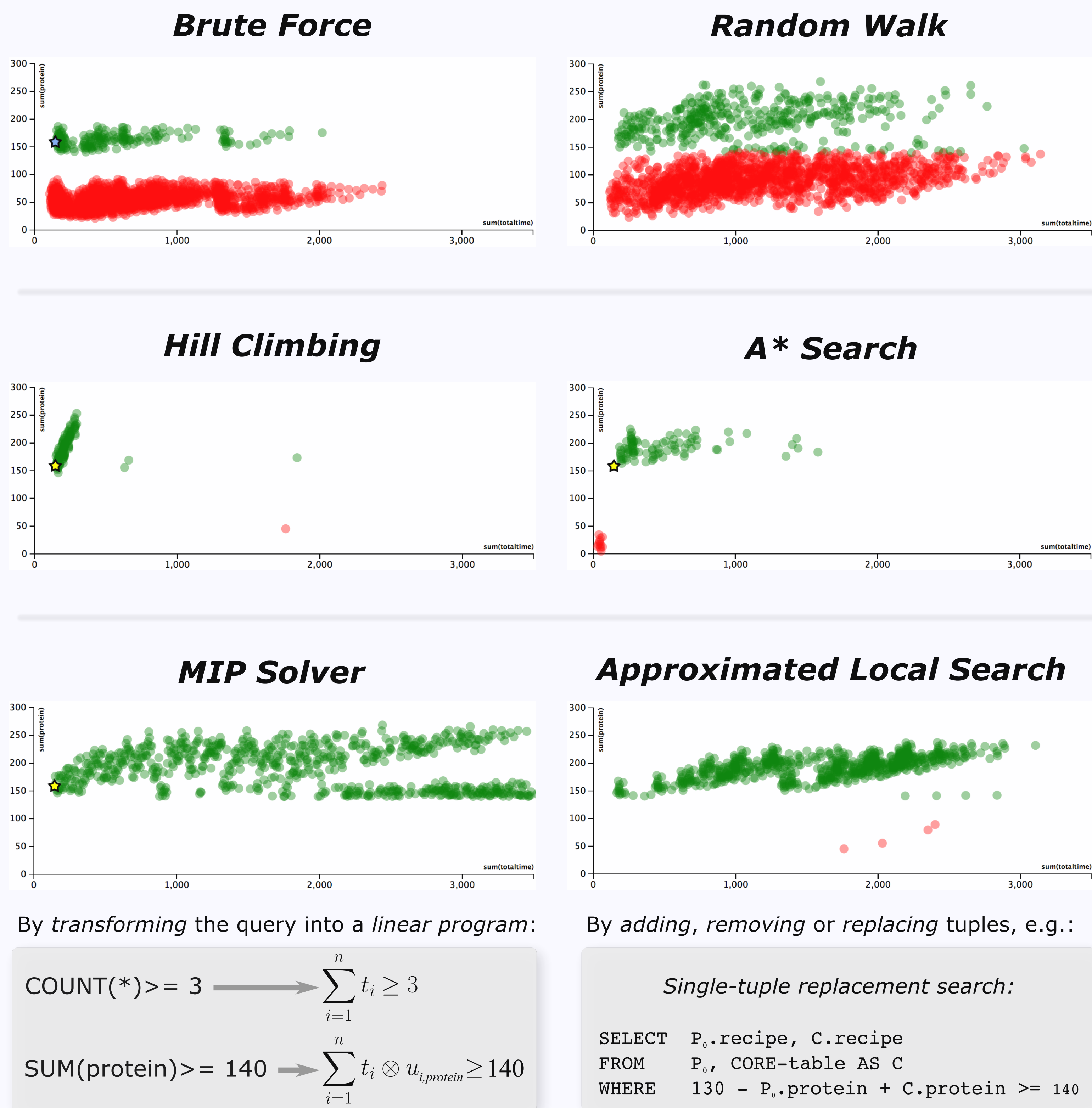
Recipe	Calorie	Protein	Prep. time	Sat. fat
Easy Wakame Brown Rice	294	4.9	55	3.3
Slow Cooker Ham and Beans	324	17.8	1210	2.1
Rice and Lentil Soup in a Jar	132	9.8	50	1.1
Boiled Crawfish	724	128.1	80	1.7

A *package* is a collection of tuples with certain *global properties* defined on the collection as a whole.

Language Specification: PaQL

```
SELECT PACKAGE(R) AS P
FROM Recipes R REPEAT 0
WHERE R.saturatedfat <= 5
SUCH THAT count(P.*) BETWEEN 3 and 6
AND sum(P.protein) >= 140
MINIMIZE sum(P.totaltime)
```

Query Evaluation



Interactive Exploration

Users can easily modify constraints. An auto-suggest feature helps with syntax

QUERY

```
SELECT PACKAGE(R)
FROM recipes_50 R
WHERE saturatedfat <= 5
SUCH THAT
  count(*) >= 3
  AND count(*) <= 6
  AND sum(protein) >= 140
MINIMIZE sum(totaltime);
```

Method: LP
Cardin. pruning: On
Display: All packages
Show relaxations: Yes
Linear problem: Show

PACKAGE

Total packages: 319

Recipe	Calorie	Carbs	Protein	Prep. time	Fat	Sat. fat	Sugar
Slow Cooker Ham and Beans	324	49.8	17.8	1210	5.9	2.1	14.9
Creamy Ham and Beans	243	27.8	17.1	35	7.0	2.5	1.9
Rhinelanders Lentil Soup	221	22.7	7.6	225	11.1	2.7	2
Caribbean Chicken with Pineapple-Cilantro Rice	361	51	28.5	50	4.8	1.2	11.7
Crawfish Pistolettes	221	18.2	13	50	10.7	4.5	4.2
Boiled Crawfish	724	30.2	128.1	80	10.1	1.7	10.5

Rows 1-6 of 6
Columns

Adaptive Exploration

Click cells or columns to get suggestions for constraints

Base Constraints

- saturatedfat ≤ 5

Global Constraints

- count(*) ≥ 3
- count(*) ≤ 6
- sum(protein) ≥ 140

SUGGESTIONS

- calories ≤ 724.0
- calories ≥ 724.0
- sum(calories) ≤ 2094.0
- sum(calories) ≥ 2094.0

Visual Summary

Packages represented: 319. Search: running. Wall-clock time: 9.55s. Optimal package found after: 0 candidate(s).

X axis: sum(totaltime) Y axis: sum(saturatedfat)

The current package's position in the result space is highlighted

Only packages found so far are visualized. Running indicates incomplete result space

The optimal package and the relaxed packages are represented with special symbols