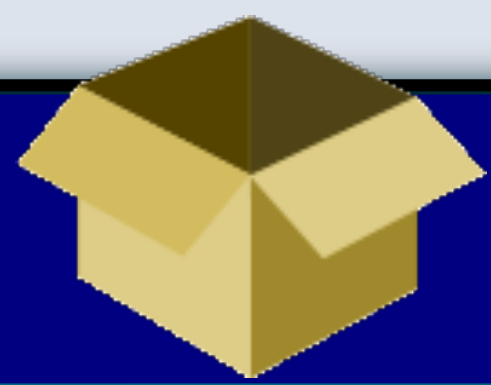




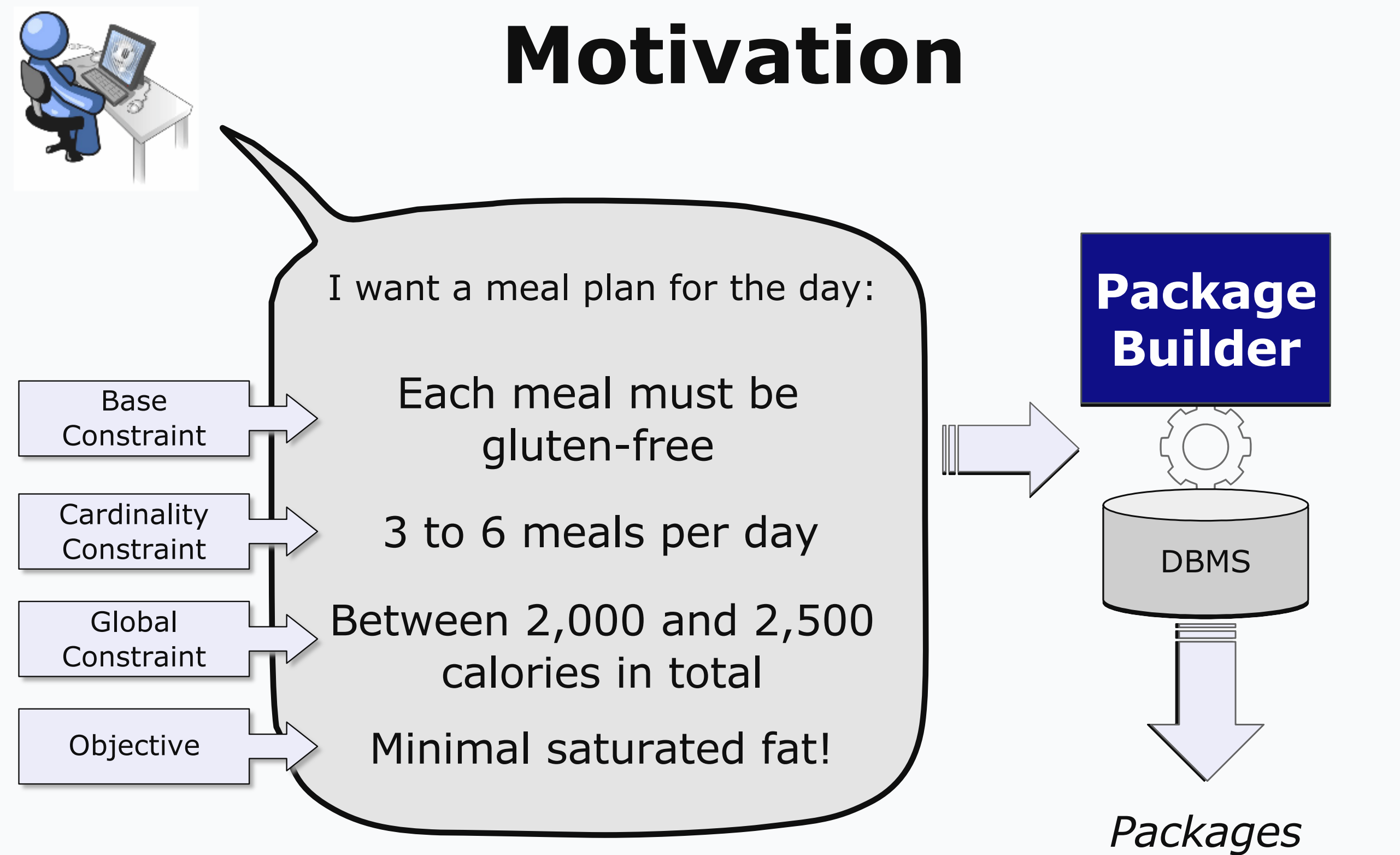
Scalable Package Queries in Relational Databases

Matteo Brucato*, Juan F. Beltran[§], Azza Abouzied[§], Alexandra Meliou*

[†]UMass Amherst [§]New York University, Abu Dhabi
http://packagebuilder.cs.umass.edu



Motivation



Packages



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

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

Language Specification: PaQL

```

SELECT      PACKAGE(R) AS P
FROM        Recipes R REPEAT 0
WHERE       R.gluten = 'free'
SUCH THAT count(P.*) BETWEEN 3 and 6
              AND sum(P.kcal) BETWEEN 2 AND 2.5
MINIMIZE  sum(P.saturated_fat)
    
```

Direct Query Evaluation

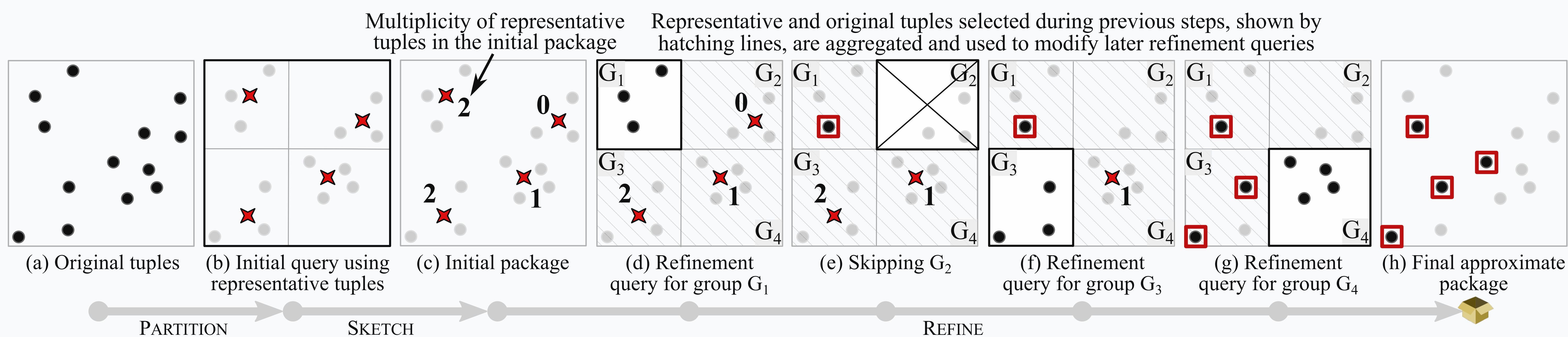
Translation: PaQL to Integer Linear Programming

$$\text{COUNT}(\ast) \geq 3 \longrightarrow \sum_{i=1}^n x_i \geq 3$$

$$\text{SUM}(\text{kcal}) \geq 2 \longrightarrow \sum_{i=1}^n x_i \times t_{i,\text{kcal}} \geq 2$$

Scalable Query Evaluation

The SketchRefine Algorithm:



Experimental Evaluation

