Relational Kernel-based Grasping with Numerical Features

Laura Antanas, Plinio Moreno, Luc De Raedt

Overview

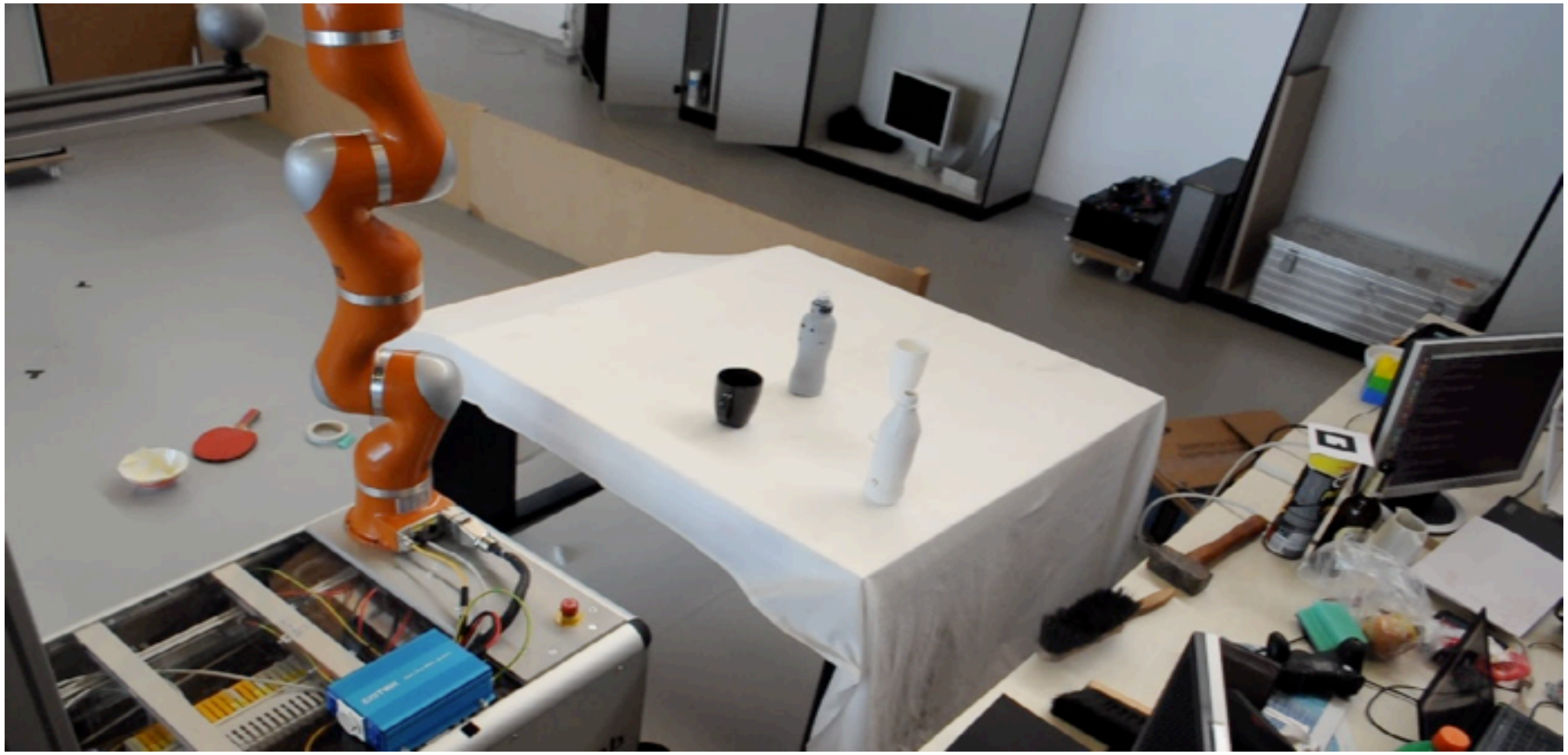
- The robot grasping task
- Motivation and contributions
- Relational problem formulation for robot grasping
- Experiments and results
- Conclusions

Overview

- **The robot grasping task**
- Motivation and contribution
- Relational problem formulation for robot grasping
- Experiments and results
- Conclusions

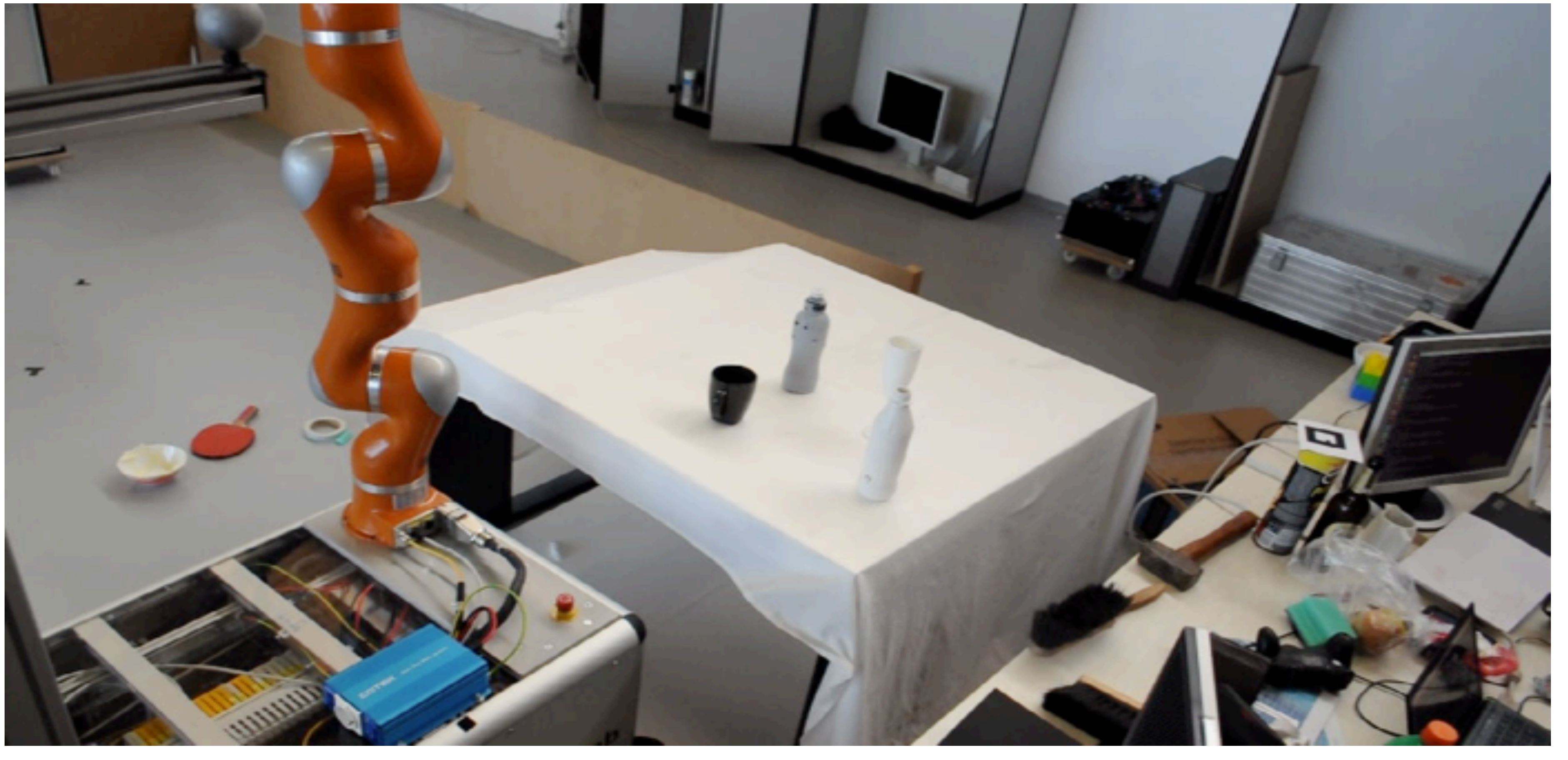
The Robot Grasping Task

Grasping scenario



The Robot Grasping Task

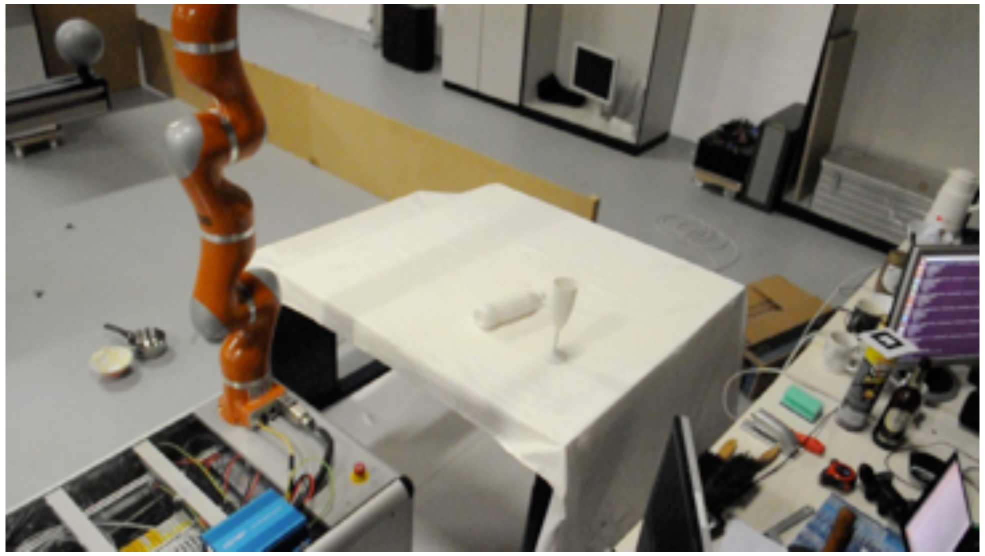
Grasping scenario



The Robot Grasping Task

Object:pan
Part:handle
Task:P&P on table

Object:bottle
Part:top
Task:pass



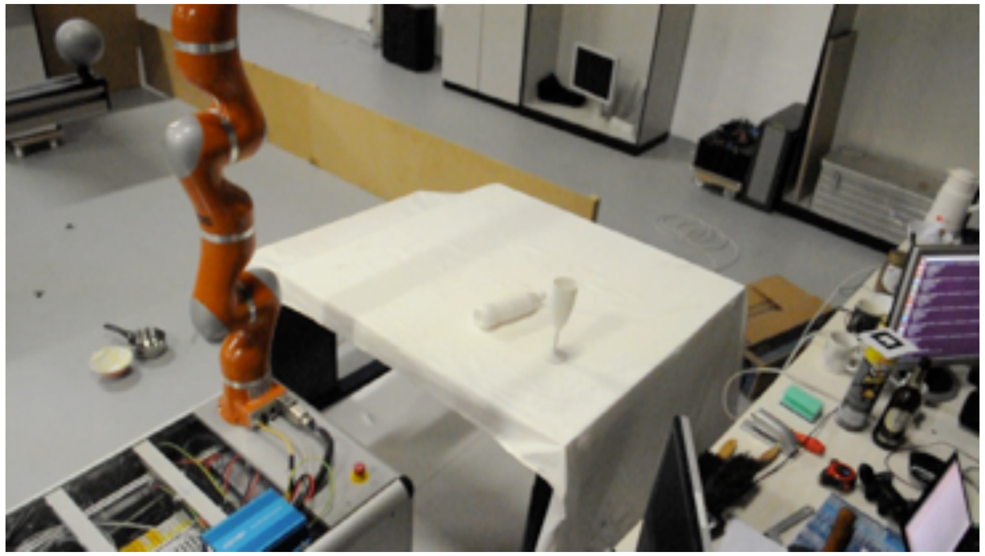
Object:bottle
Part:middle
Task:Pass

Object:cup
Part:top
Task:Pour out

The Robot Grasping Task

Object:pan
Part:handle
Task:P&P on table

Object:bottle
Part:top
Task:pass



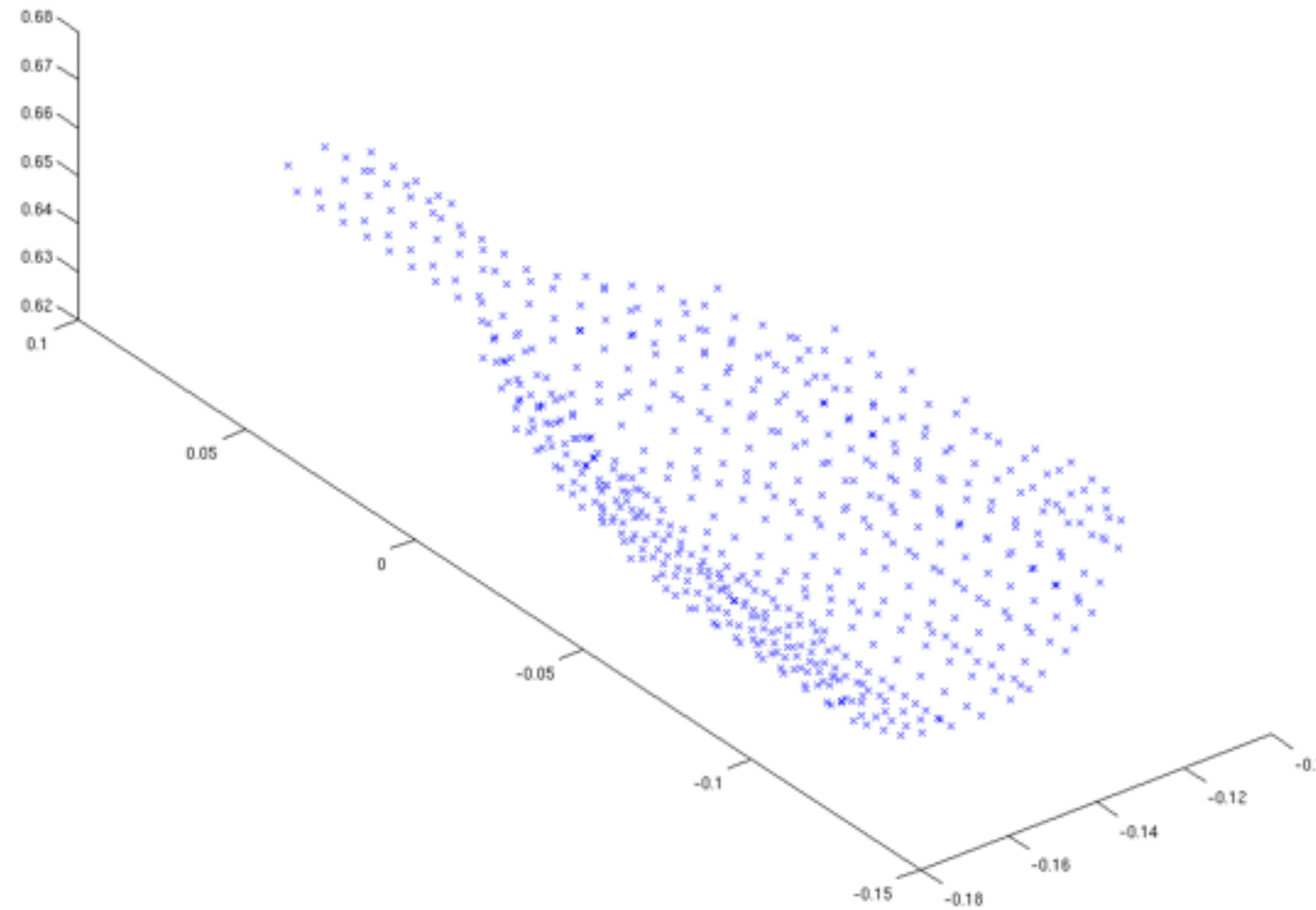
Object:bottle
Part:middle
Task:Pass

Object:cup
Part:top
Task:Pour out

The Robot Grasping Task

Find the pre-grasp pose, that is where to place the gripper with respect to the object, in order to execute a stable grasp

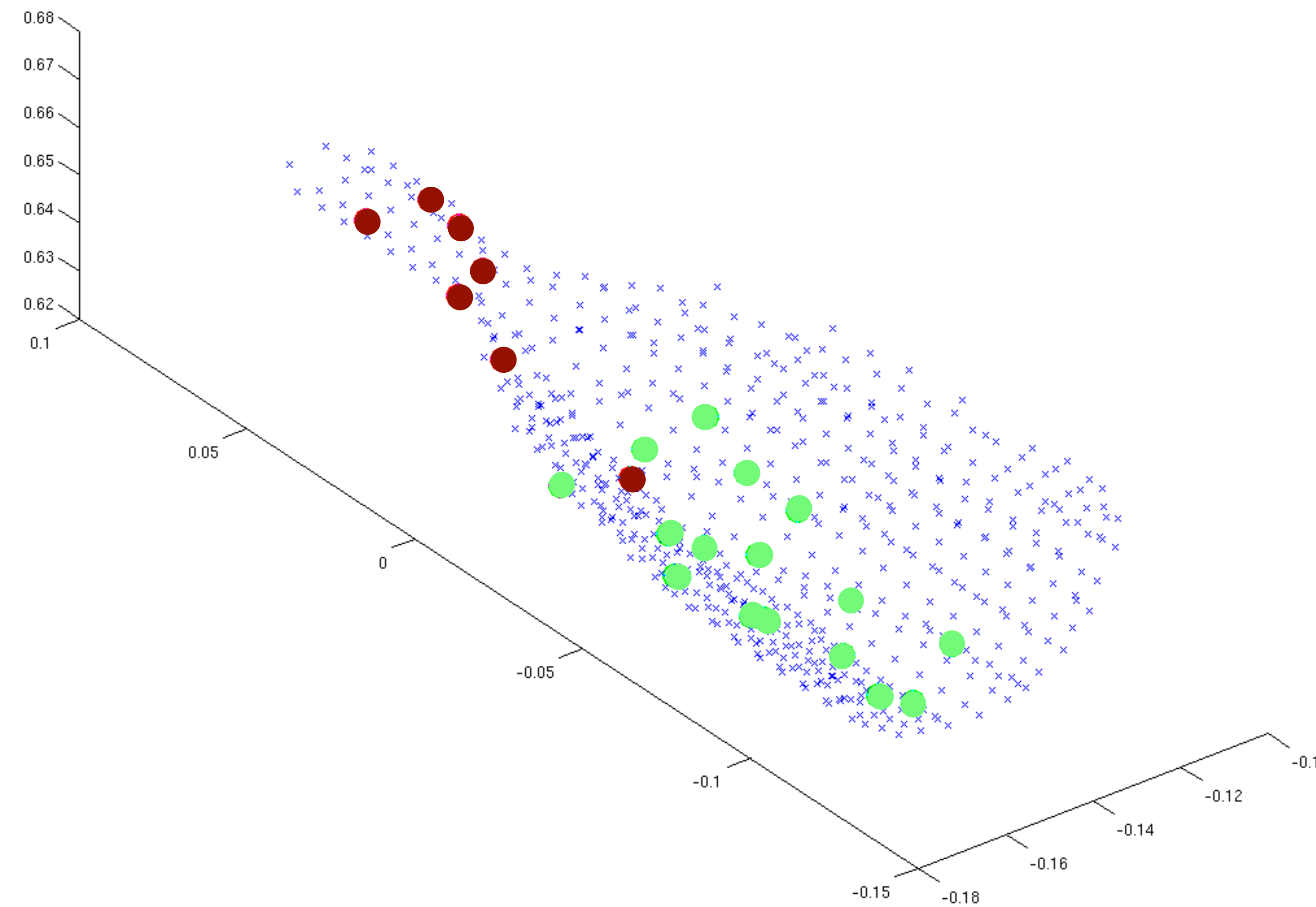
Task: recognize graspable points of an object



The Robot Grasping Task

Find the pre-grasp pose, that is where to place the gripper with respect to the object, in order to execute a stable grasp

Task: recognize graspable points of an object

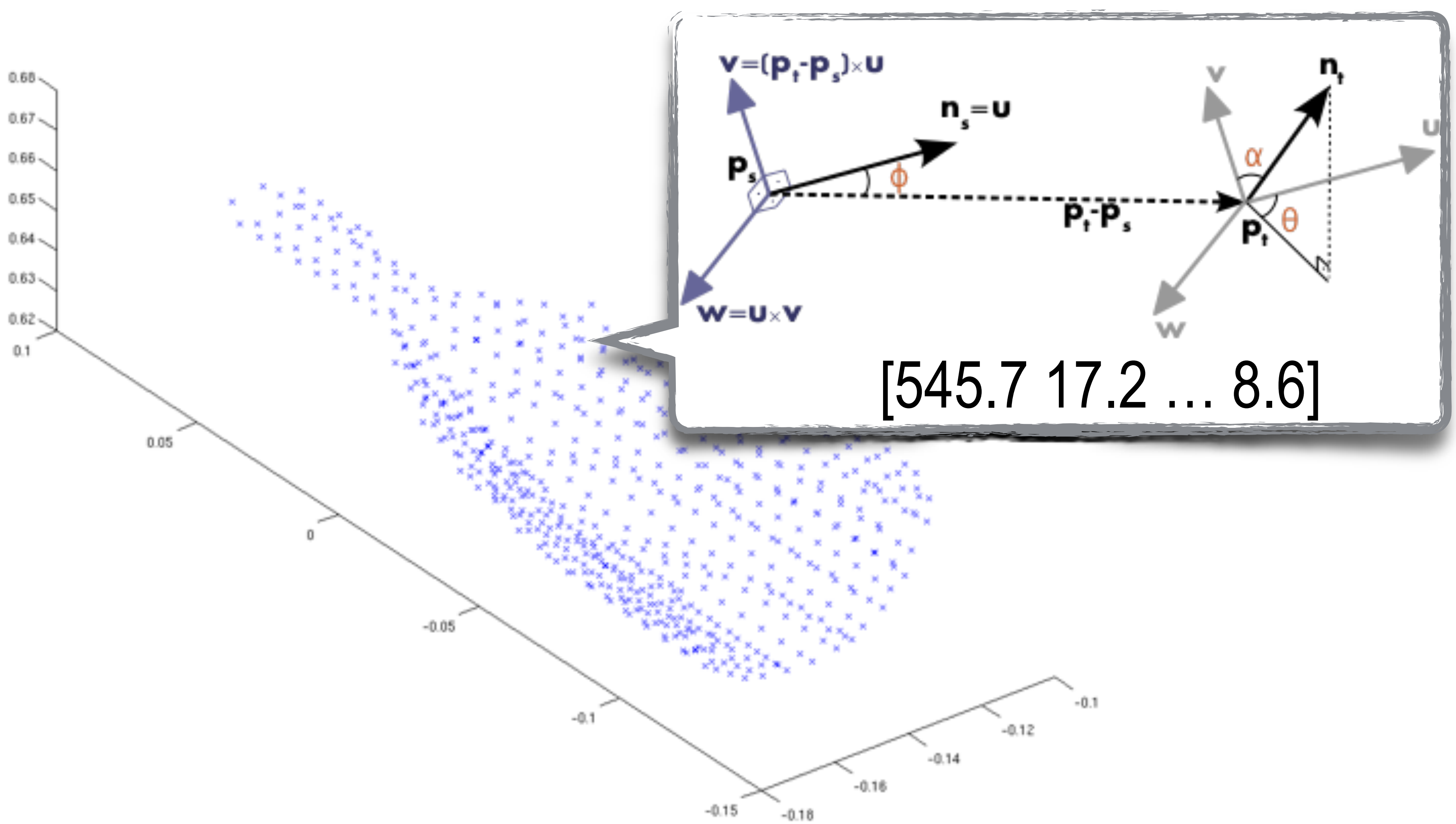


Overview

- The robot grasping task
- **Motivation and contribution**
- Relational problem formulation for robot grasping
- Experiments and results
- Conclusions

Motivation and Contribution

Local visual descriptors to characterise the graspability of an object point: e.g., point feature histogram (PFH), VFH, 3D SC



Motivation and Contribution

However, such local shape features do not work properly on more complex or (self-) occluded objects.

Motivation and Contribution

ILP Point of view

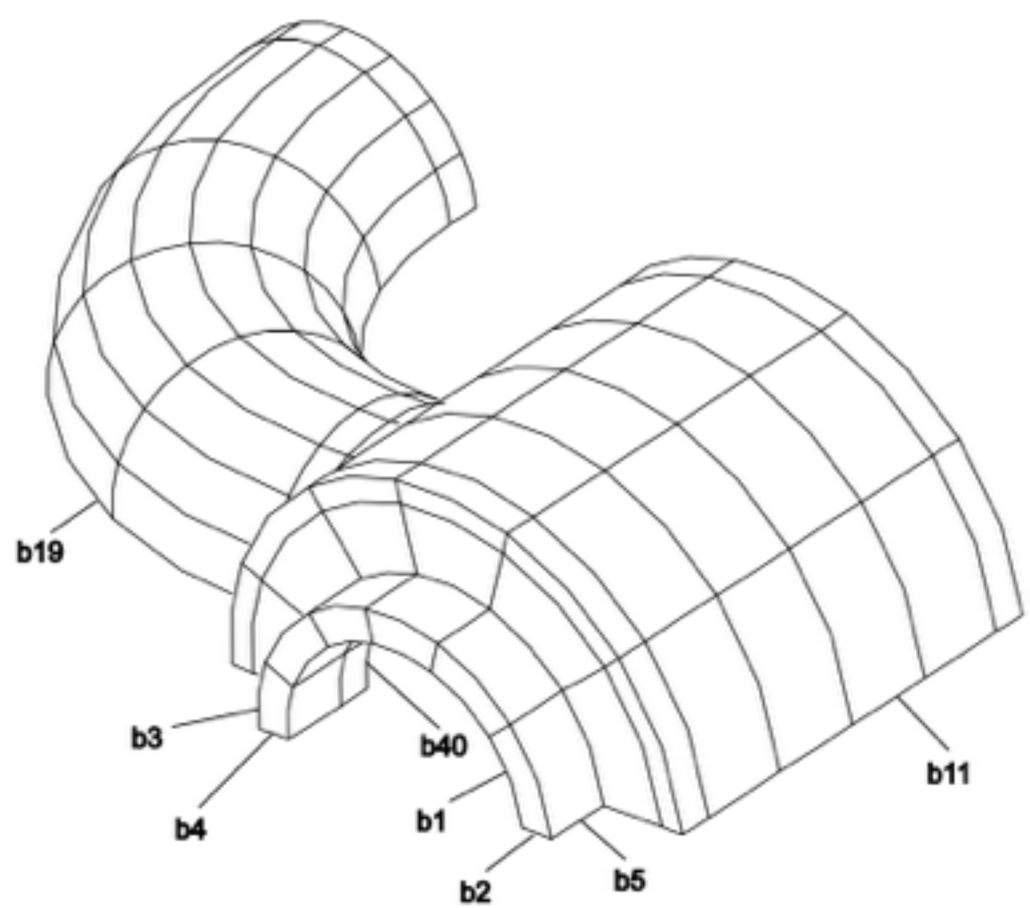


Figure 12.1: A typical structure with its corresponding FE mesh. From Dolšak [1991].

Motivation and Contribution

ILP Point of view

- Meshes :
  - finite / discrete — purely qualitative
  - structure and background relations matter
- Point clouds :
  - continuous — purely quantitative
  - structure and background relations matter
  - understandability less important

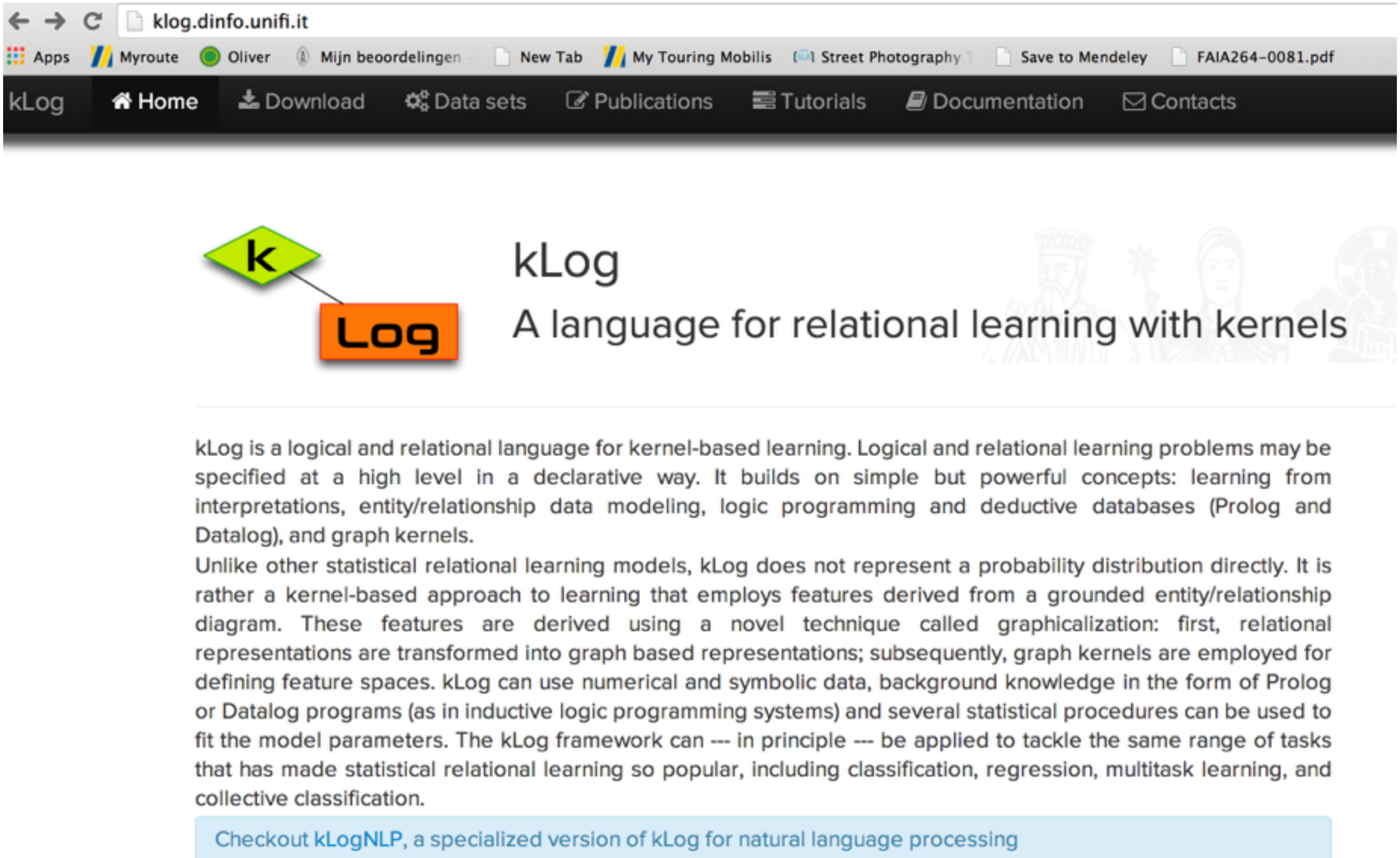
Motivation and Contribution

Contributions

- we show that the extended structure of the object can improve robot grasping:
  - using SRL, we build a graph-based representation of the object exploiting both local numerical features and higher-level information about the structure of the object — extended contextual shape information of the object.
- we contribute a relational kernel-based approach to numerical feature pooling for robot grasping:
  - for each descriptor of the object point, our relational kernel exploits extended contextual information by pooling numerical shape features

Overview

- The robot grasping task
- Motivation and contribution
- **Relational problem formulation for robot grasping**
- Experiments and results
- Conclusions



klog.dinfo.unifi.it

Apps | Myroute | Oliver | Mijn beoordelingen | New Tab | My Touring Mobilis | Street Photography | Save to Mendeley | FAIA264-0081.pdf

kLog | Home | Download | Data sets | Publications | Tutorials | Documentation | Contacts

k Log

kLog

A language for relational learning with kernels

kLog is a logical and relational language for kernel-based learning. Logical and relational learning problems may be specified at a high level in a declarative way. It builds on simple but powerful concepts: learning from interpretations, entity/relationship data modeling, logic programming and deductive databases (Prolog and Datalog), and graph kernels.

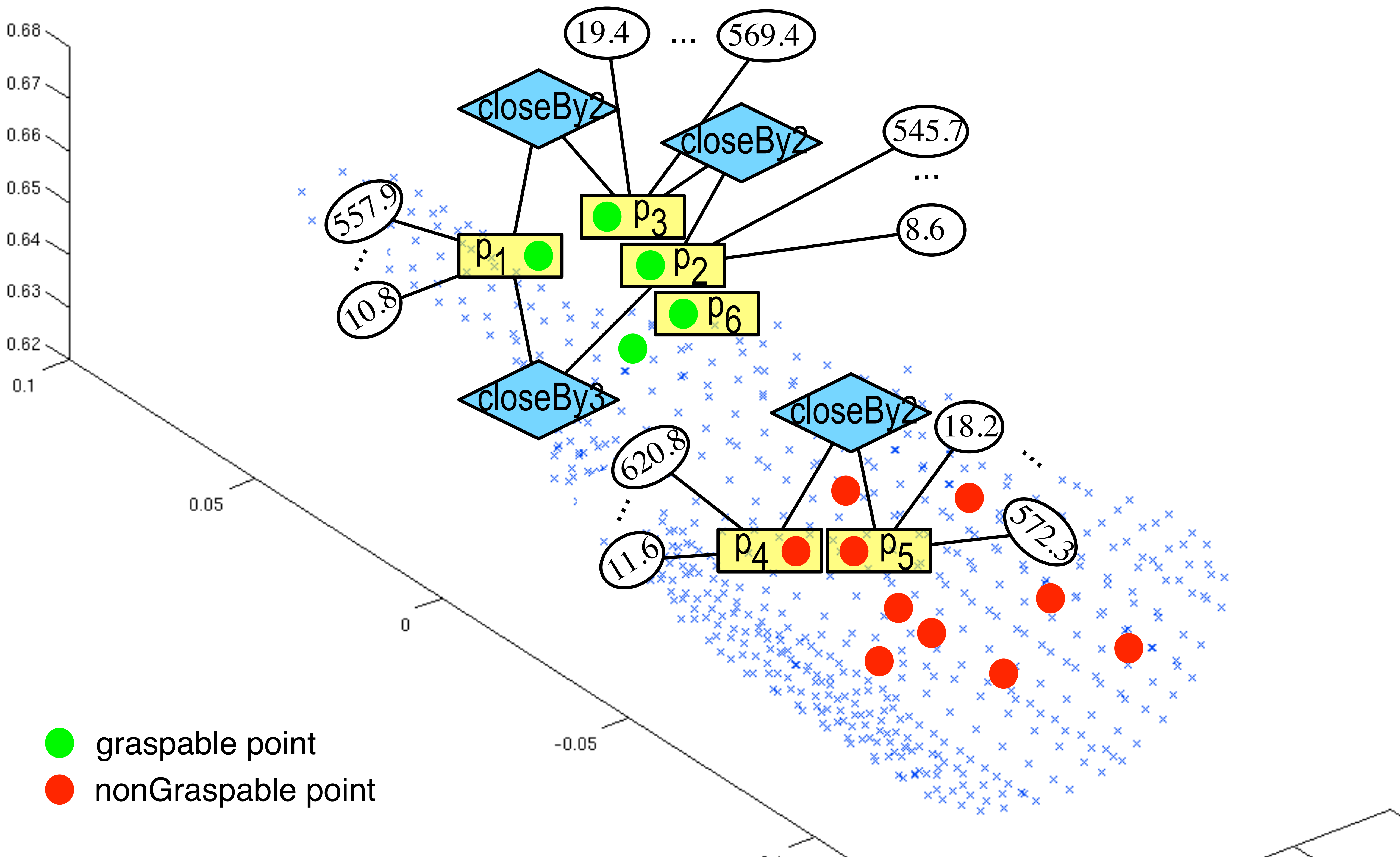
Unlike other statistical relational learning models, kLog does not represent a probability distribution directly. It is rather a kernel-based approach to learning that employs features derived from a grounded entity/relationship diagram. These features are derived using a novel technique called graphicalization: first, relational representations are transformed into graph based representations; subsequently, graph kernels are employed for defining feature spaces. kLog can use numerical and symbolic data, background knowledge in the form of Prolog or Datalog programs (as in inductive logic programming systems) and several statistical procedures can be used to fit the model parameters. The kLog framework can --- in principle --- be applied to tackle the same range of tasks that has made statistical relational learning so popular, including classification, regression, multitask learning, and collective classification.

Checkout kLogNLP, a specialized version of kLog for natural language processing

Frasconi, Costa, De Raedt, De Grave, AIJ 14

Verbeke, Frasconi, Costa, De Raedt, De Grave, ACL-Demo, 14

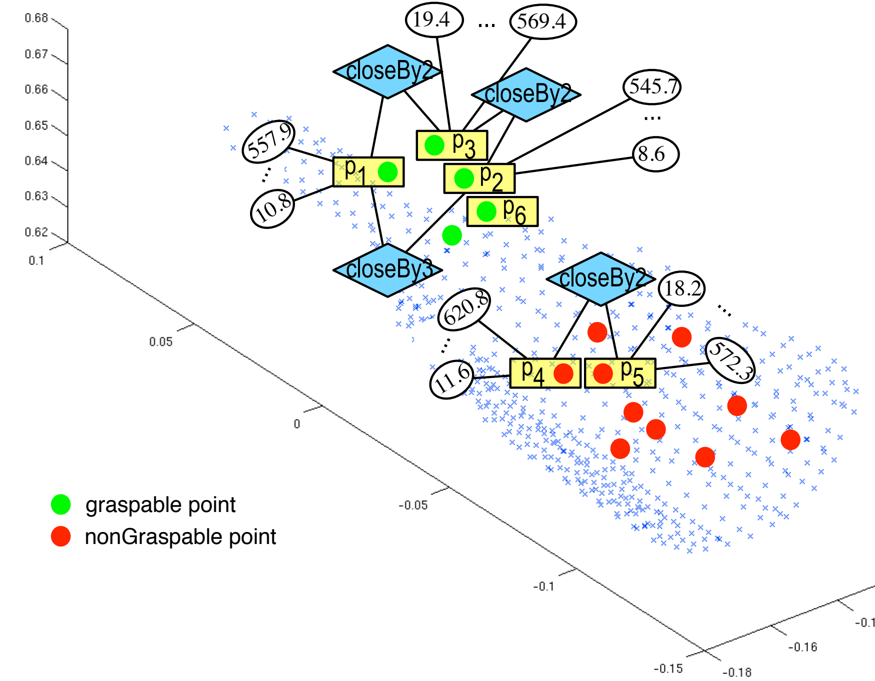
Relational Robot Grasping



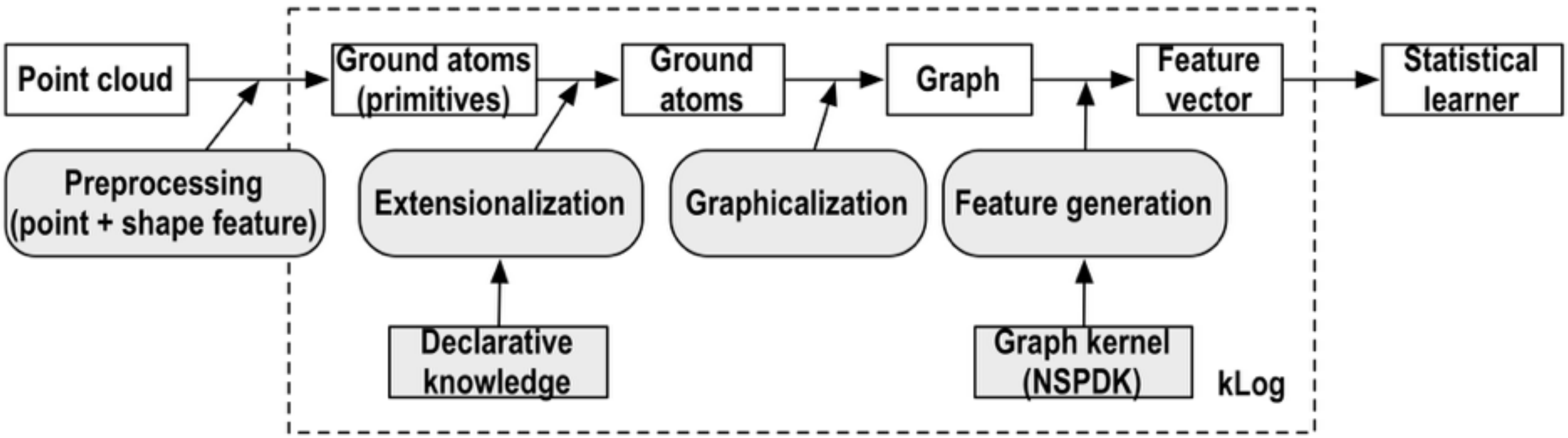
graspable point

nonGraspable point

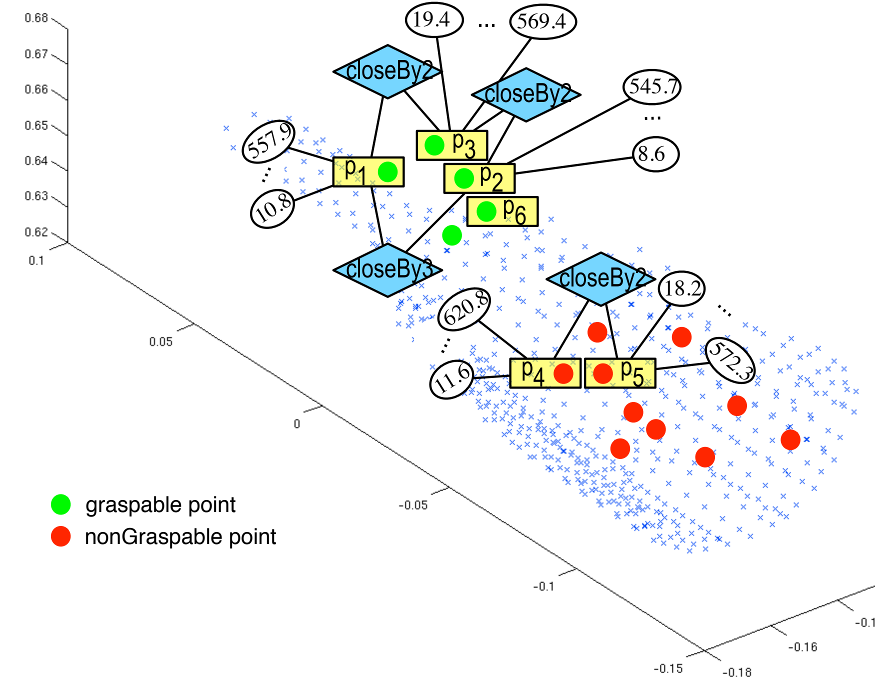
Relational Robot Grasping



The SRL framework: *kLog* [Frasconi et al., '12]



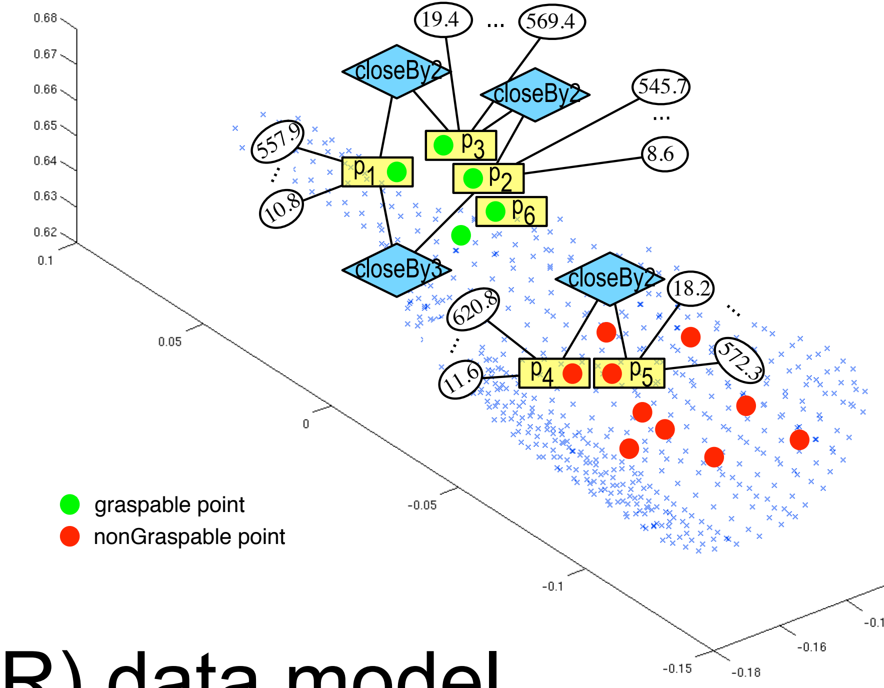
Relational Robot Grasping



Grasping primitives

- grasping *reaching points*
- their *3D locations*
- their numerical *shape features: 3D shape context*, *point feature histogram* (PFH)*, viewpoint feature histogram* (VFH)

Relational Robot Grasping



Data modelling: classic entity/relationship (E/R) data model

- entity sets

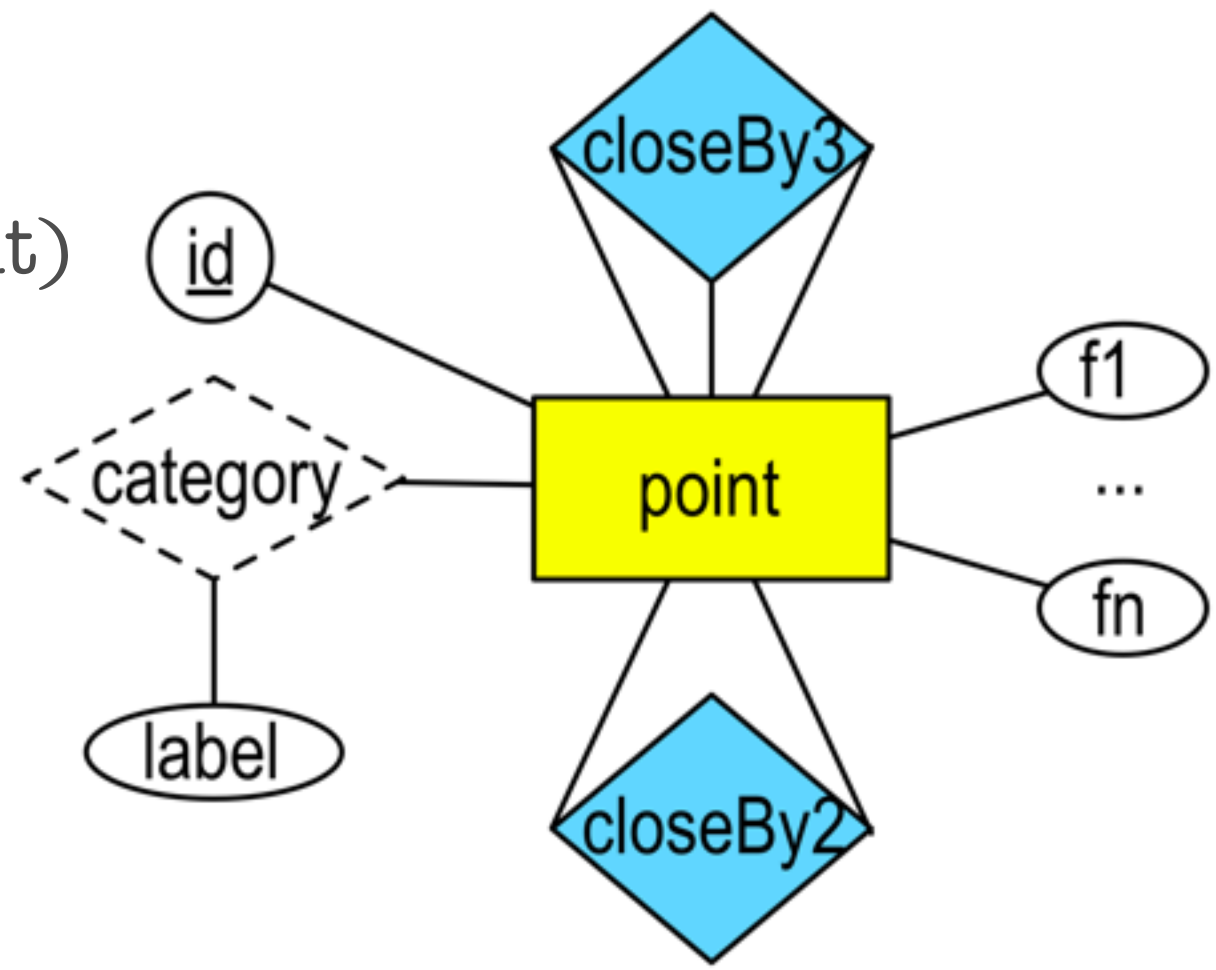
  point(id::self, f1::property, ..., fn::property)

- relationships

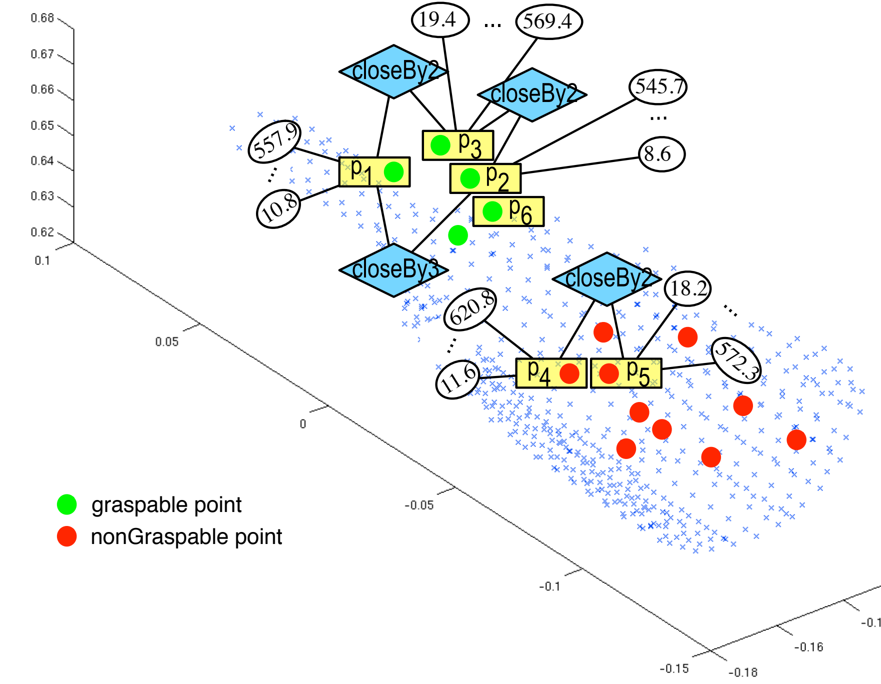
  closeBy2 (id_1::point, id_2::point)

- attributes

  f1, ..., fn, label



Relational Robot Grasping



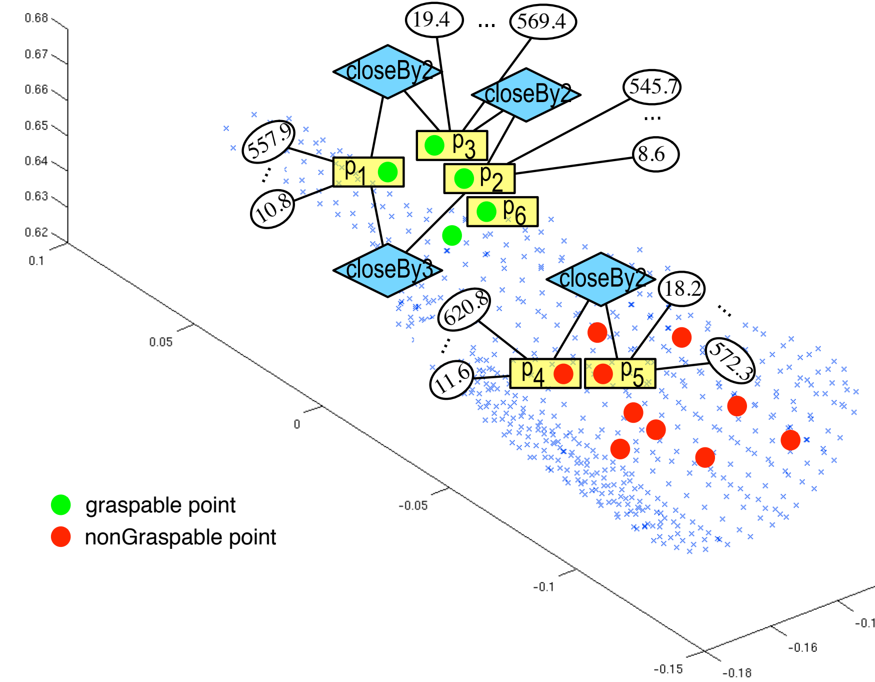
R-relations as background knowledge

- declarative feature construction

$\mathbf{closeBy2}(P1,P2) \leftarrow point(P_1,F_{11},\ldots,F_{1n}), point(P_2,F_{21},\ldots,F_{2n}), sameCloud(P1,P2), edist(P1,P2,Dist), Dist < T.$

- T is a constant calculated for every object as a ratio relative to the object dimension

Relational Robot Grasping



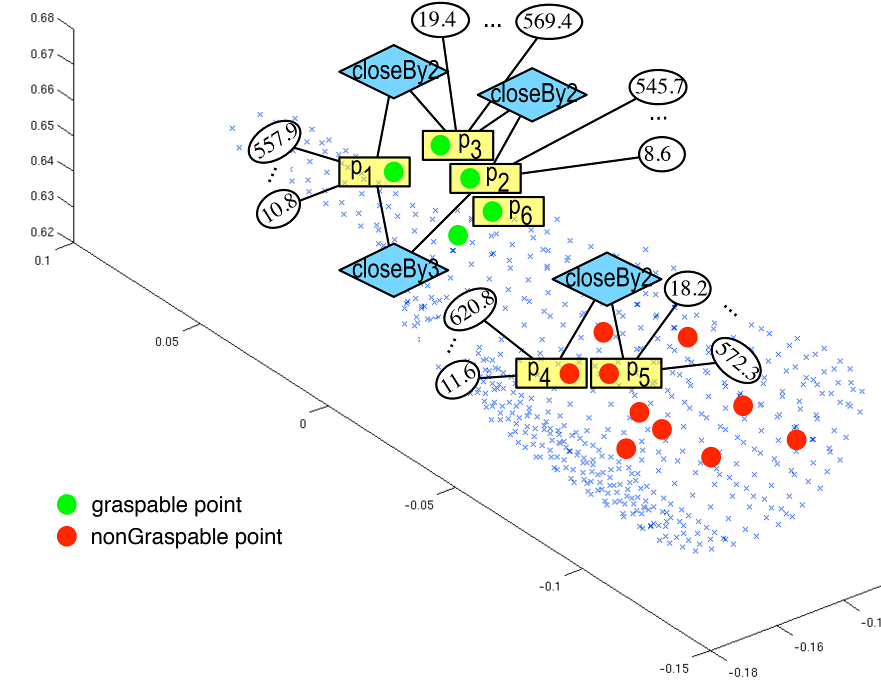
Point cloud interpretation

- each point cloud is represented as an instance of a relational database (i.e., as a set of relations)

x ={**point**(p_1, 10.8, … , 557.9), **point**(p_2, 8.6, … , 545.7), **point**(p_3, 19.4, … , 569.4), **point**(p_4, 11.6, … , 620.8), … , **closeBy2**(p_1,p_3), **closeBy2**(p_3,p_2), **closeBy2**(p_4,p_5), …, **closeBy3**(p_1,p_2,p_3), …}.

y = {**category**(p_1,nonGrasp), **category**(p_2,nonGrasp), **category**(p_3,nonGrasp), **category**(p_4,grasp), …}.

Relational Robot Grasping



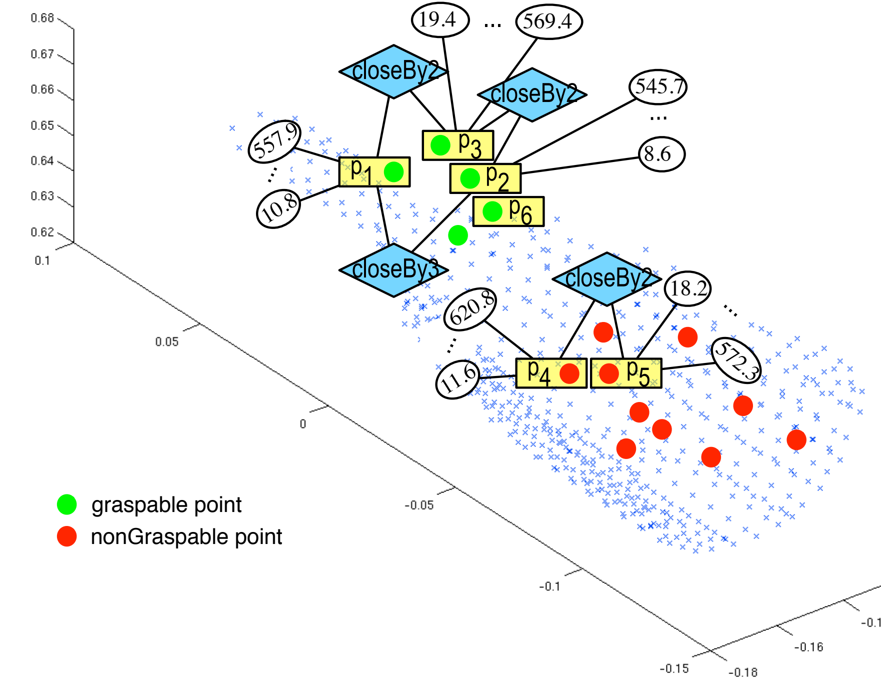
Point cloud interpretation

- each point cloud is represented as an instance of a relational database (i.e., as a set of relations)

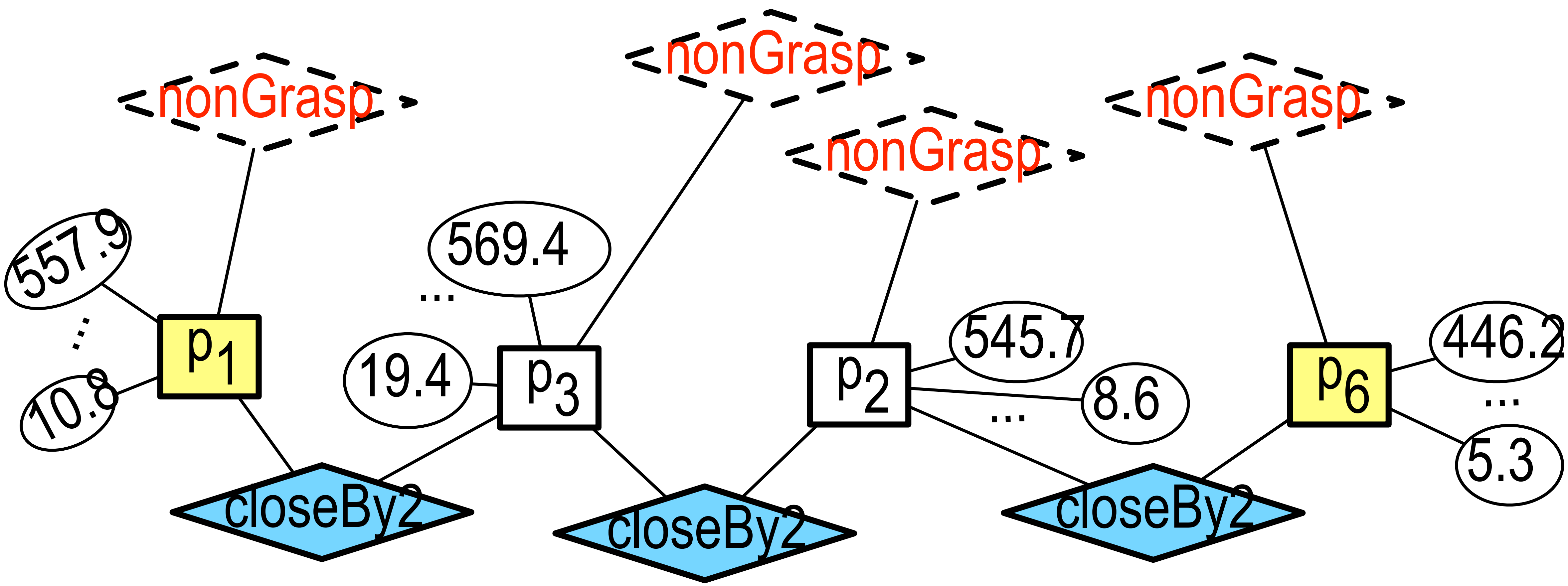
x ={**point**(p_1, 10.8, … , 557.9), **point**(p_2, 8.6, … , 545.7), **point**(p_3, 19.4, … , 569.4), **point**(p_4, 11.6, … , 620.8), … , **closeBy2**(p_1,p_3), **closeBy2**(p_3,p_2), **closeBy2**(p_4,p_5), …, **closeBy3**(p_1,p_2,p_3), …}.

y = {**category**(p_1,?), **category**(p_2,?), **category**(p_3,?), **category**(p_4,?), …}.

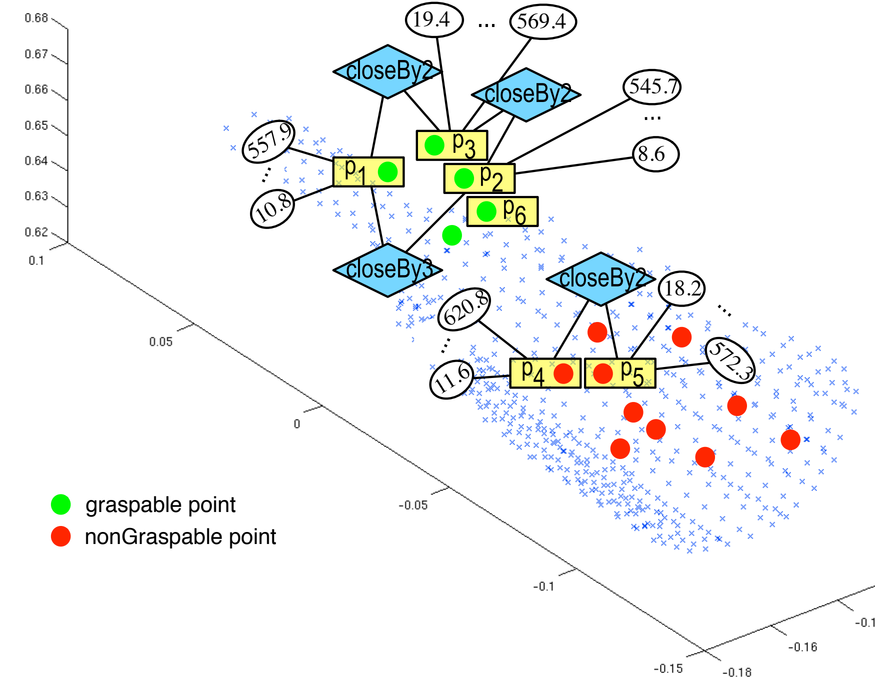
Relational Robot Grasping



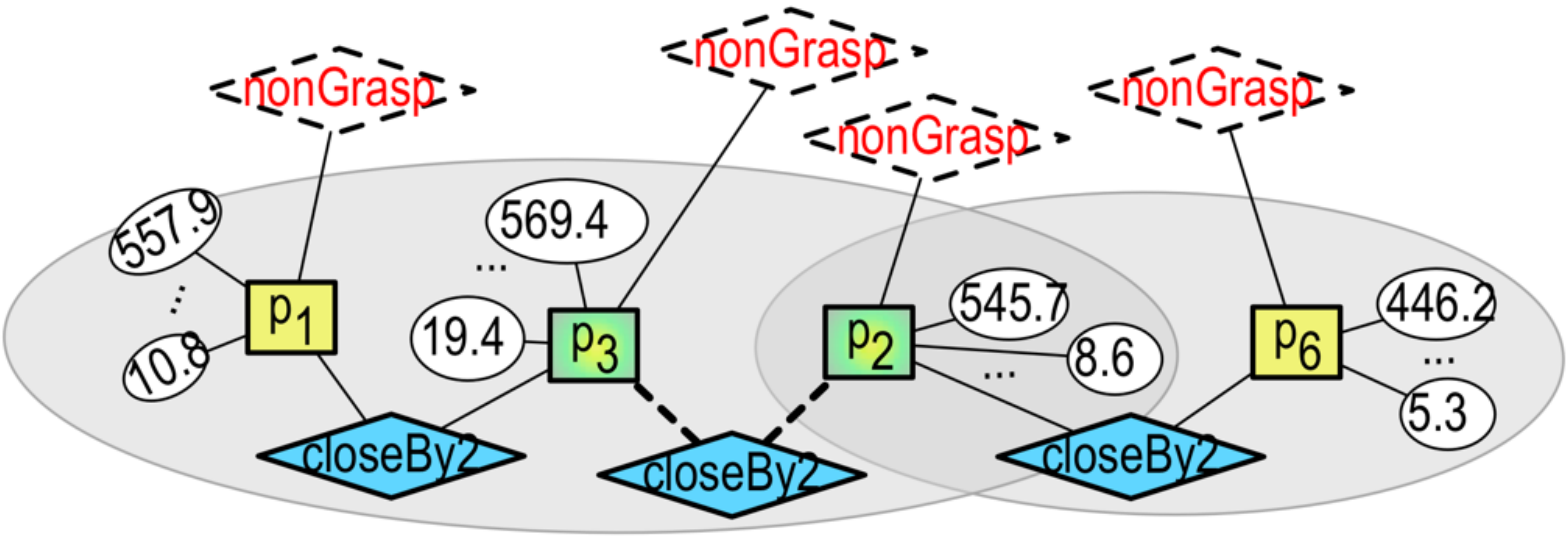
Graphicalization



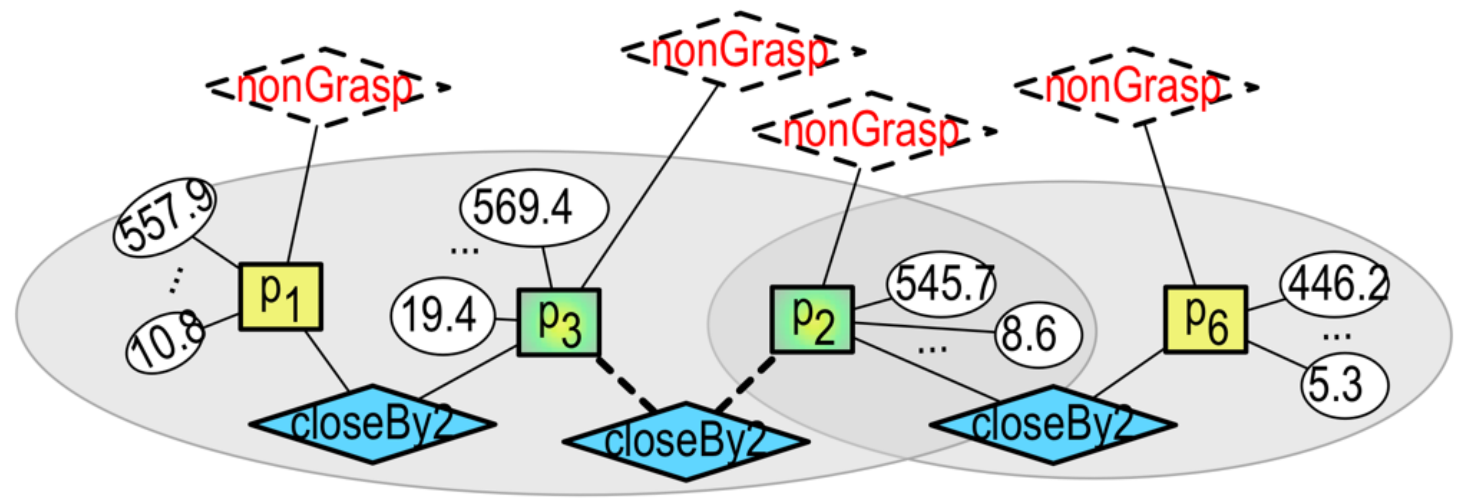
Relational Robot Grasping



Feature Generation



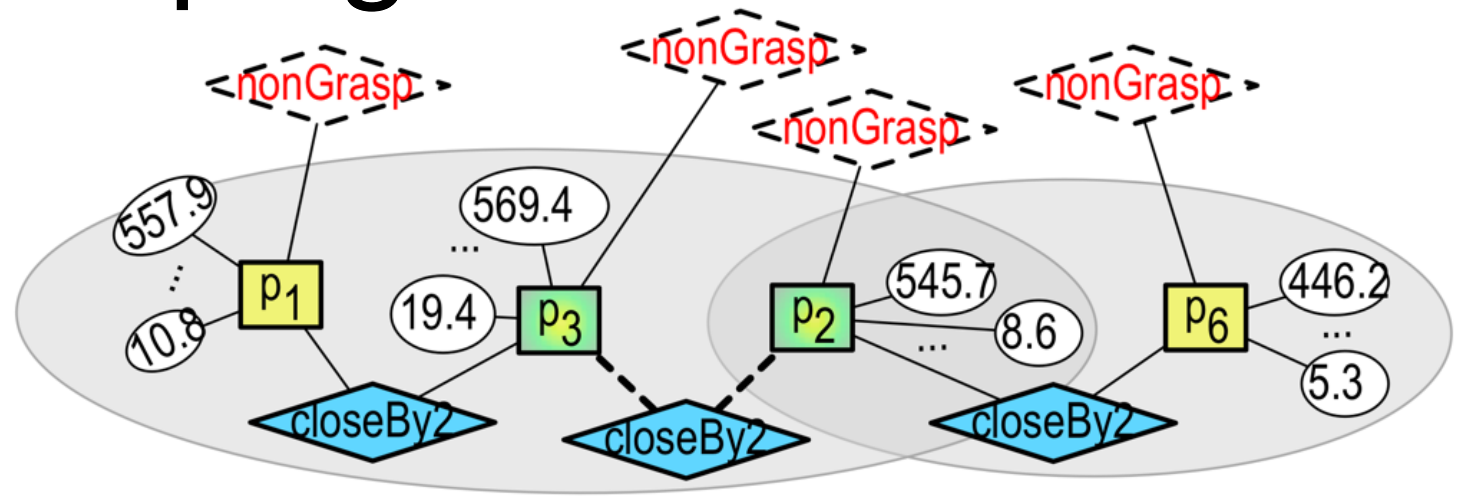
Relational Robot Grasping



Feature Generation

- decomposition kernel between two graphs [Costa and De Grave,'10] counting the number of common parts
- $R_{r,d} = \{(N_r^v(G), N_r^u(G), G): d^*(u,v)=d\}$, $r=0,\ldots,R$, $d=0,\ldots,D$
- $N_r^v(G)$ = subgraph A rooted in v with radius r
- $N_r^u(G)$ = subgraph B rooted in u with radius r
- A,B: $R_{r,d}^{-1}(A,B,G)$

Relational Robot Grasping



The decomposition kernel is defined by relations $R_{r,d}$:

$$K(G,G') = \sum_{r=0}^{R} \sum_{d=0}^{D} \sum_{\substack{A,B:\, R_{r,d}^{-1}(A,B,G) \\ A',B':\, R_{r,d}^{-1}(A',B',G')}} k((A,B),(A',B')).$$

- $k((A,B),(A',B')) = 1$ iff (A,B) and (A',B') are pairs of isomorphic subgraphs — hard match kernel
- $k((A,B),(A',B'))$: multinomial distribution of labels in (A,B) or (A',B') — soft match kernel
- hard match on discrete labels & soft match on numerical labels — hard-soft match kernel

Now : we would use Orsini's GIKs [IJCAI 15]

Soft Match

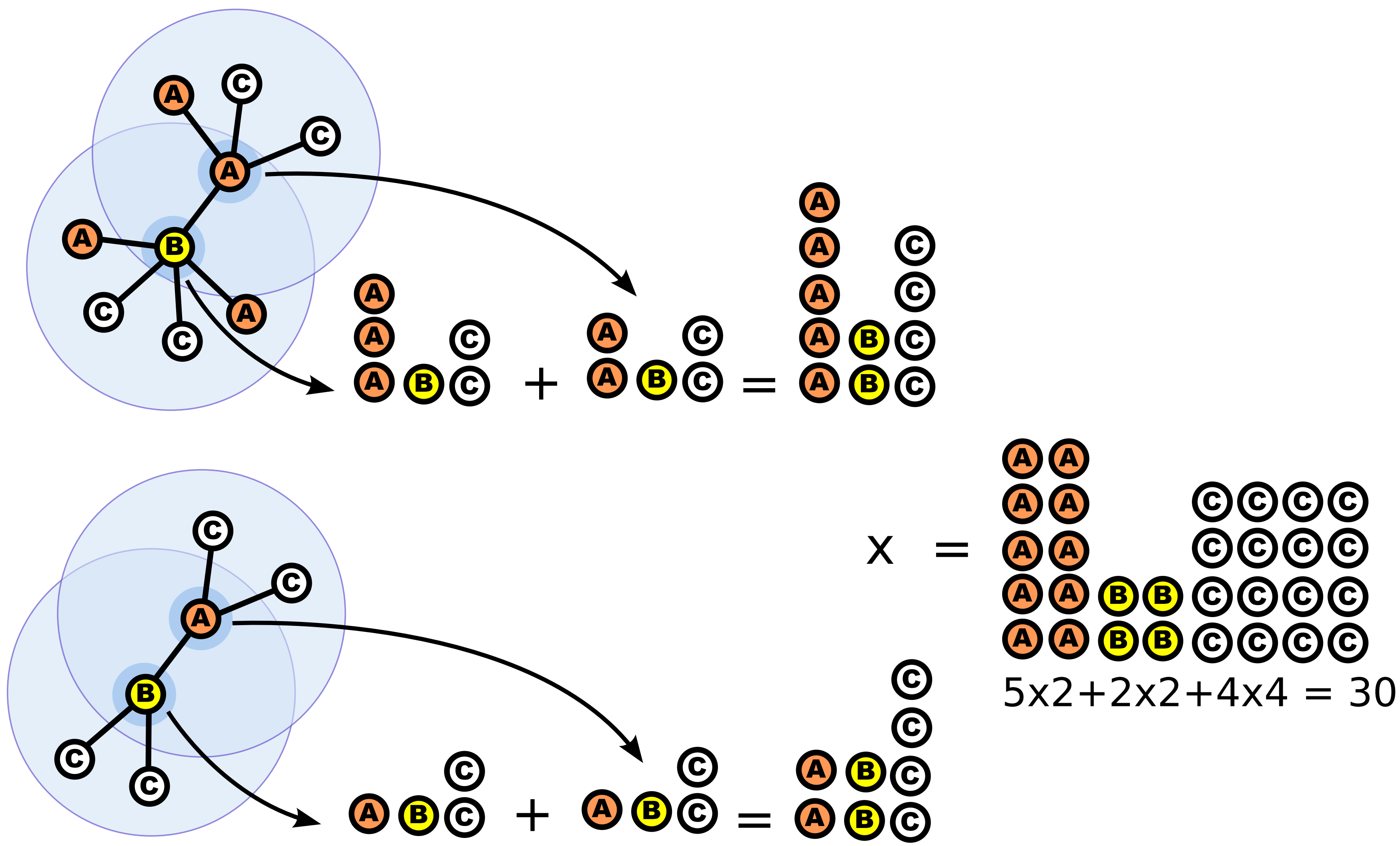


Figure 4: Illustration of the soft matching kernel. Only features generated by a selected pair of vertices are represented: vertices A and B at distance 1 yield a multinomial distribution of the vertex labels in neighborhoods of radius 1. On the right we compute the contribution to the kernel value by the represented features.

Overview

- The robot grasping task
- Motivation and contribution
- Relational problem formulation for robot grasping
- **Experiments and results**
- Conclusions

Experiments and Results

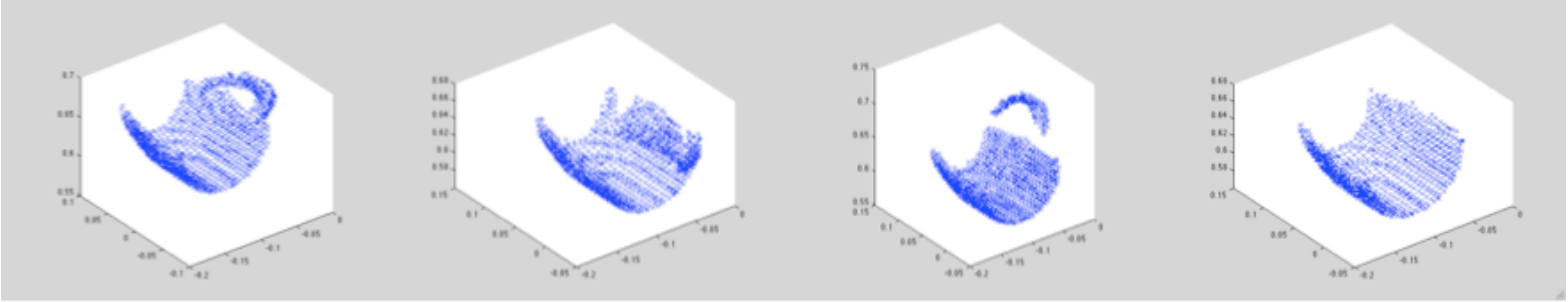
Q_1: Does numerical shape feature pooling improve upon local shape features for the robot grasping task considered?

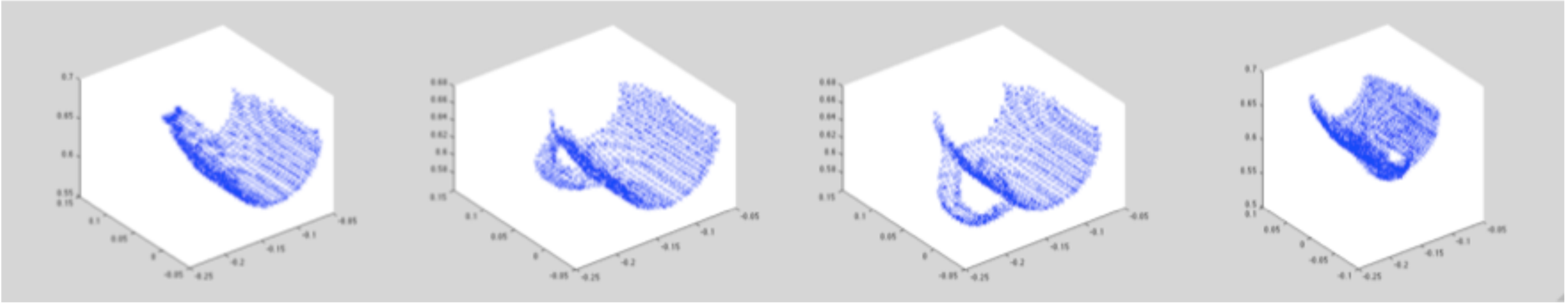
Q_2: Does hard-soft matching improve over soft matching when incorporating contextual shape information?

Experiments and Results

Dataset (robot simulator)

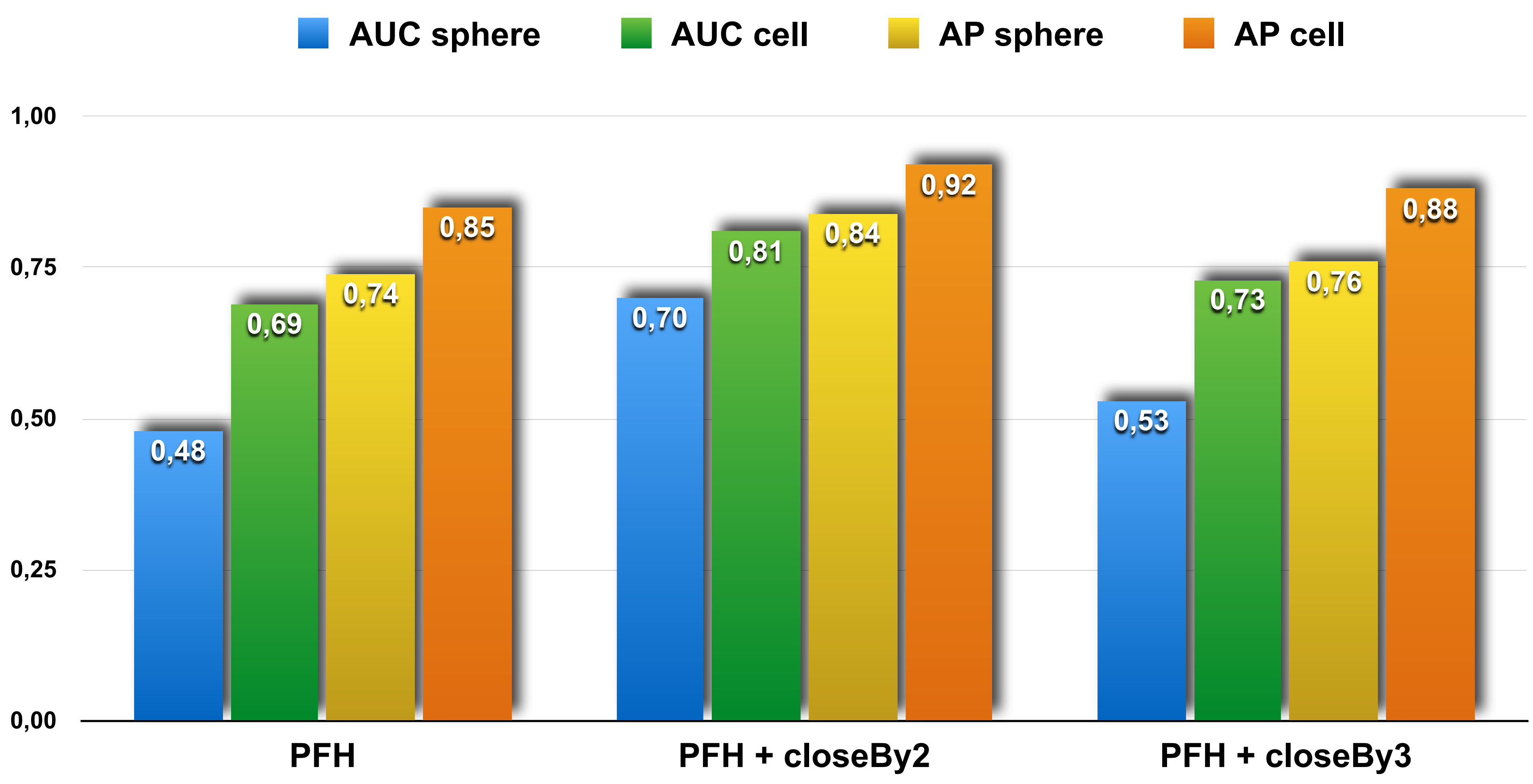
- 8 objects: ellipse, rectangle, rounded object, 2 glasses, 3 cups
- 2631 instances (1972 positives and 659 negatives)





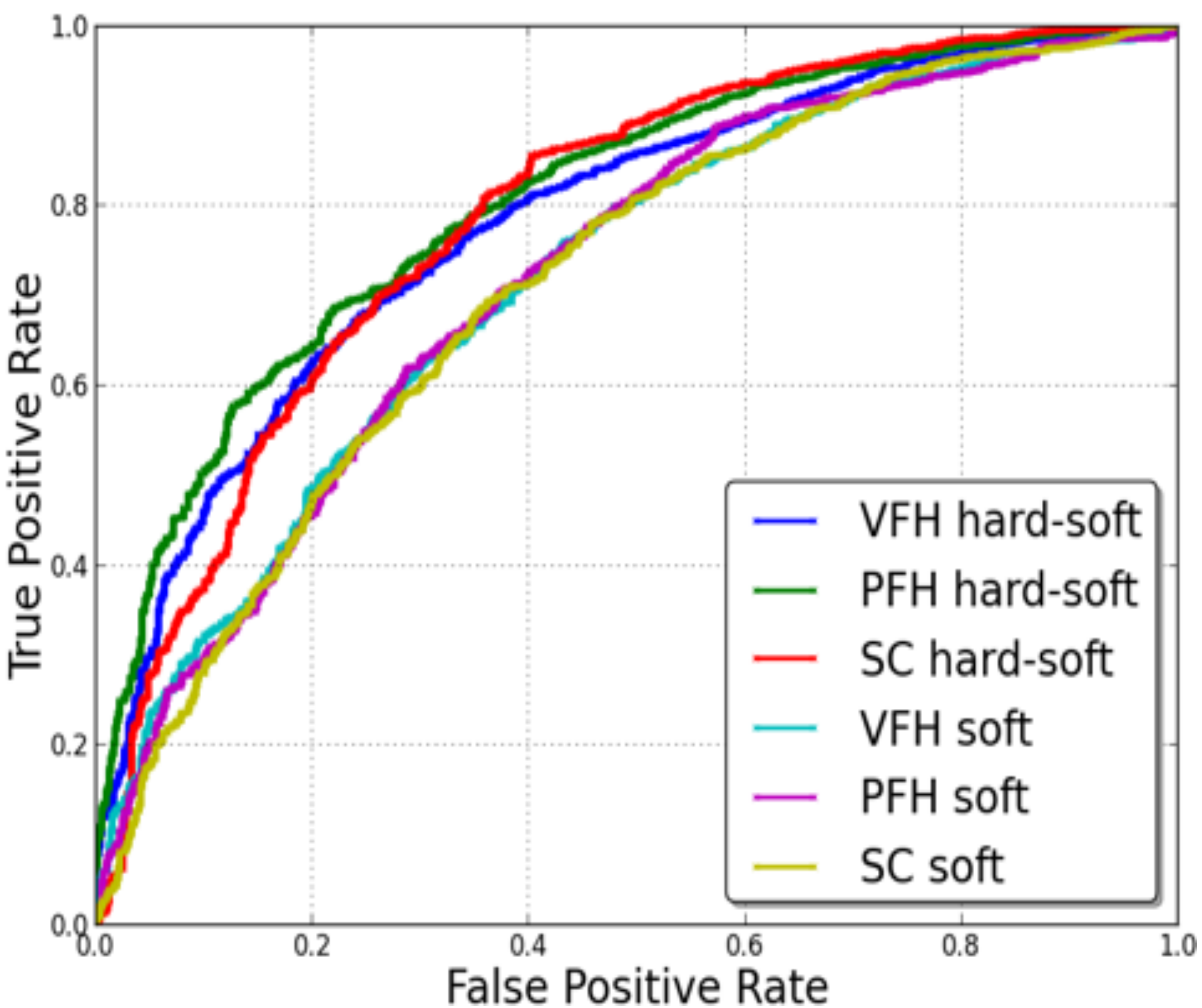
Experiments and Results

Shape feature pooling via hyper-parameters R=2,D=2 improves upon local shape features (Q_1)

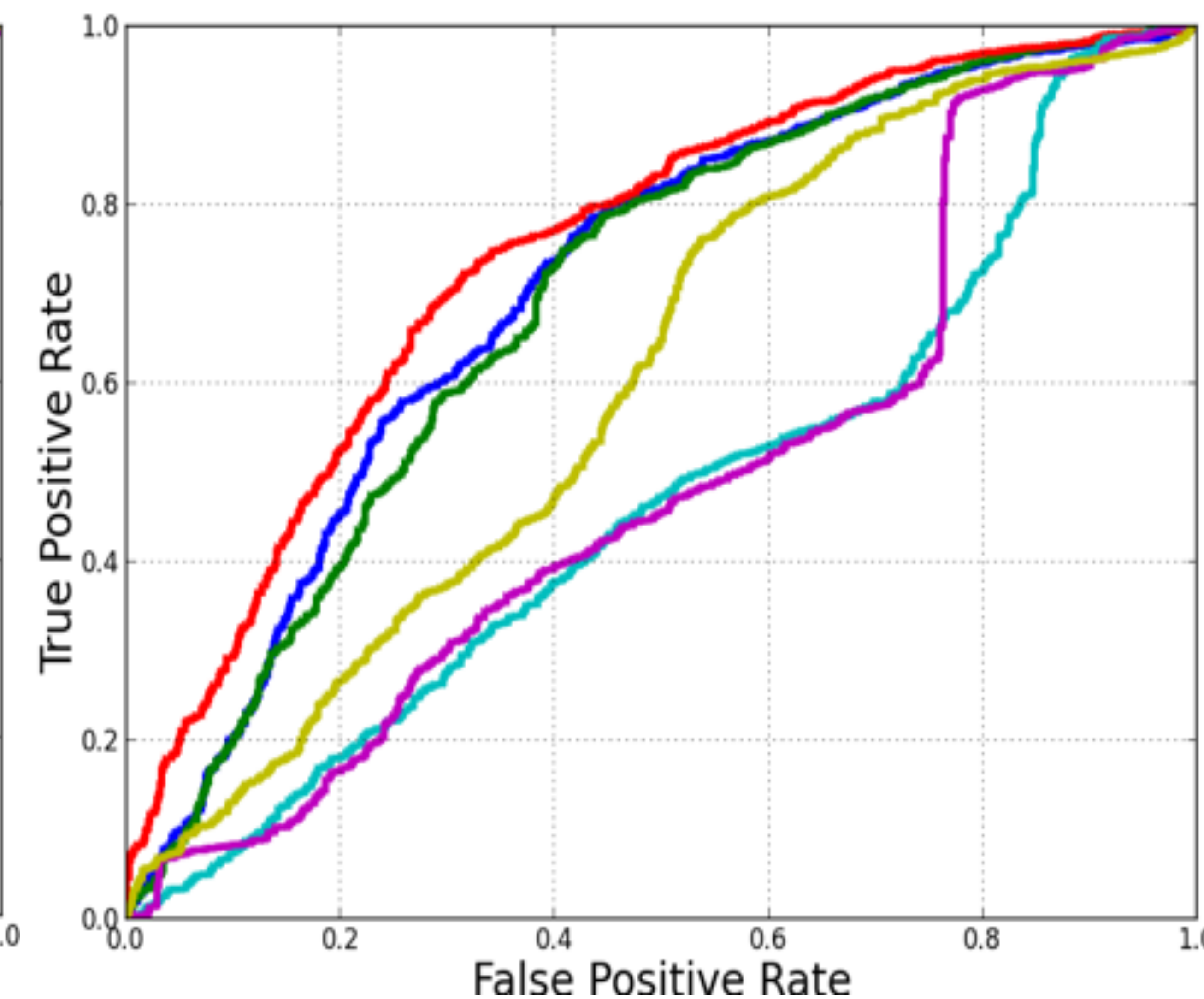


Experiments and Results

Hard-soft vs soft matching (Q_2)



grasping cell



grasping sphere

Overview

- The robot grasping task
- Motivation and contributions
- Relational problem formulation for robot grasping
- Experiments and results
- **Conclusions**

Conclusions

We propose a relational kernel-based approach to recognize graspable object points

- extended contextual object shape information is encoded via qualitative spatial relations among object points
- kernels on graphs are used to compute highly discriminative features based on contextual information

We show experimentally that pooling spatially related numerical shape feature improves robot grasping results upon purely local shape-based approaches.