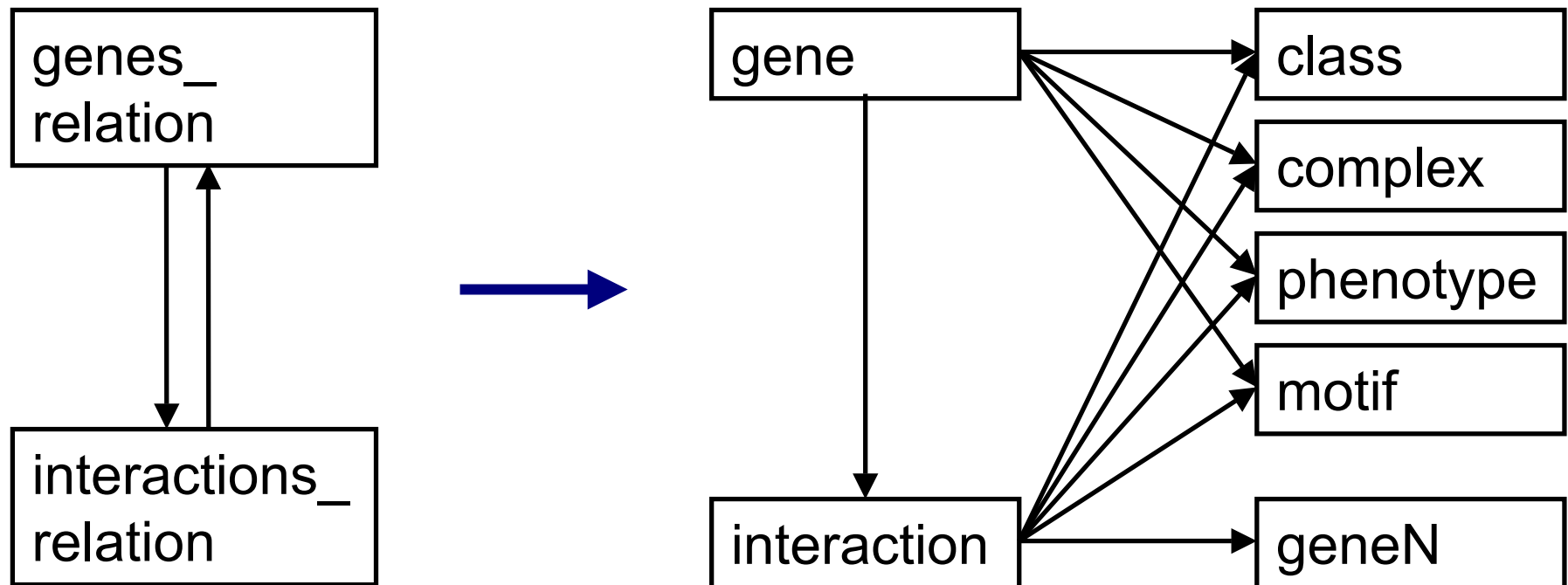# KDD Cup 2001:
Gene/Protein Function Prediction Using the Multirelational Learning Algorithm RELAGGS

Mark-A. Krogel

Otto-von-Guericke-Universität, Magdeburg, Germany
School of Computer Science, Research Group Stefan Wrobel

Preparation: A Multirelational Task

☐ General: renormalize into multiple tables as a natural representation of the data

☐ Specific for KDD Cup tasks 2/3: consider only interactions with high correlations, assume transitivity, make symmetry explicit

Algorithm: RELAGGS [Krogel/Wrobel:ILP01]

☐ Computes selected joins following user-defined *foreign links*

☐ Performs automatic transformation of multiple tables into single table with the help of aggregate functions
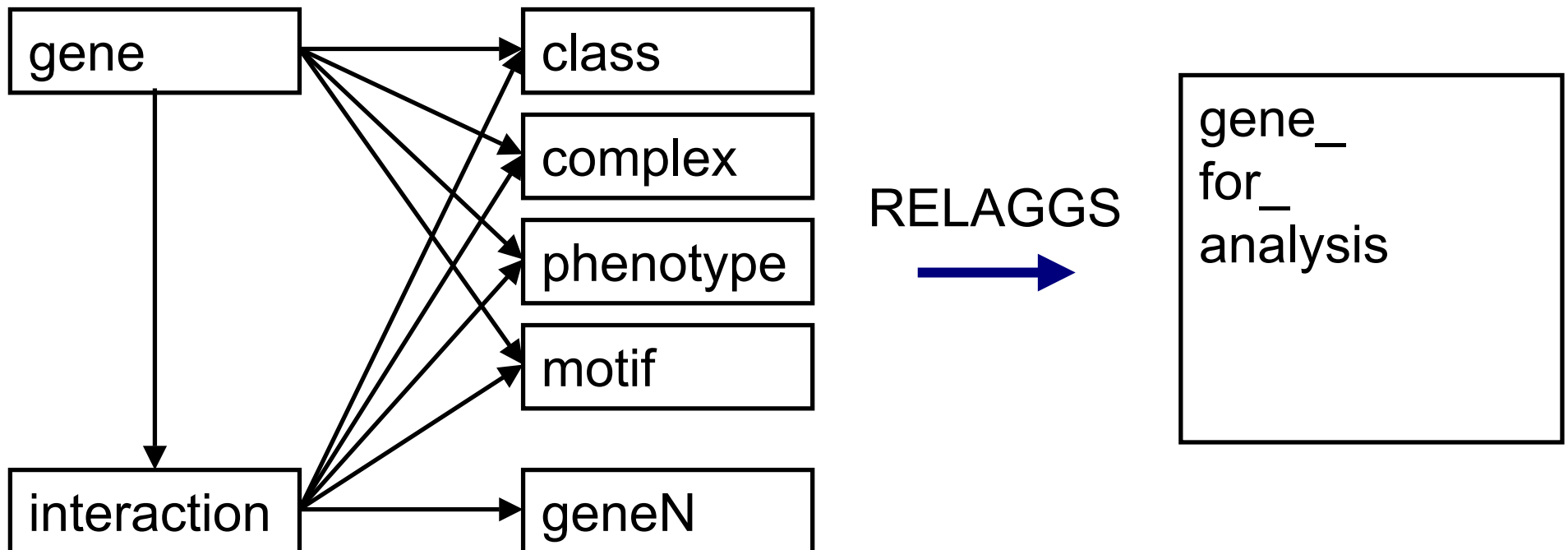
☐ Uses propositional learner such as C4.5 or SVM*light*

Summary

- □ RELAGGS allows to work with natural <u>multirelational form</u> of data immediately

- □ Easy specification of possible joins with <u>foreign links</u>

- □ Maximal preservation of information through <u>aggregation</u>

- □ Accuracies:
  93,6% on task 2: rank 1
  69,8% on task 3: rank 4

- □ http://kd.cs.uni-magdeburg.de

- □ M.-A. Krogel, S. Wrobel: Transformation-Based Learning Using Multirelational Aggregation. 11th International Conference On Inductive Logic Programming, Strasbourg, France, Sept. 2001.