Relational Markov Networks (RMNs)

Ch 6, Intro to SRL

Overview

- Classification in relational domains
- Previous approaches
- Discriminative models
- Results

The Problem

- Classification in relational domains
- Example: WebKB
- Classify web pages as:

  course, faculty, project, student, other

- How can we exploit link structure?

Flat classification

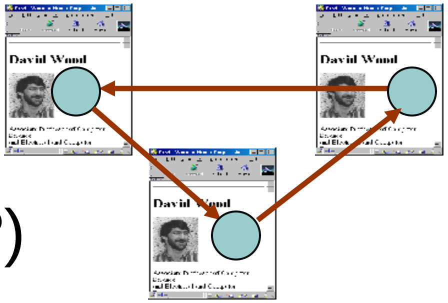
- Simplest: words on page
- Using links: anchor words
  - Problem: usually proper names
- URLs of neighboring page as features
  - Problem: does not generalize to new sites
- What we really want is:
  - Use labels of neighboring pages
  - Use structure of web pages and link graph
  - Identify and exploit sections, hubs

“Model-free” methods

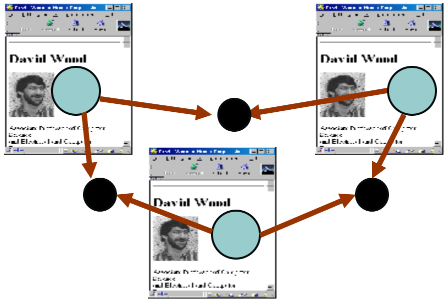
- Chakrabarti et al. (1998), Neville & Jensen (2000)
  - Iteratively relabel based on neighbors
- Slattery and Mitchell (2000)
  - FOIL + Naïve Bayes bag-of-words
- Problem:
  - No coherent global model
  - Ad hoc learning and “inference” procedures

PRMs

- Extension of Bayesian networks…
  - (Koller and Pfeffer,98)
- Want edges between linked pages
  - Problem: can’t have direct links (Why?)



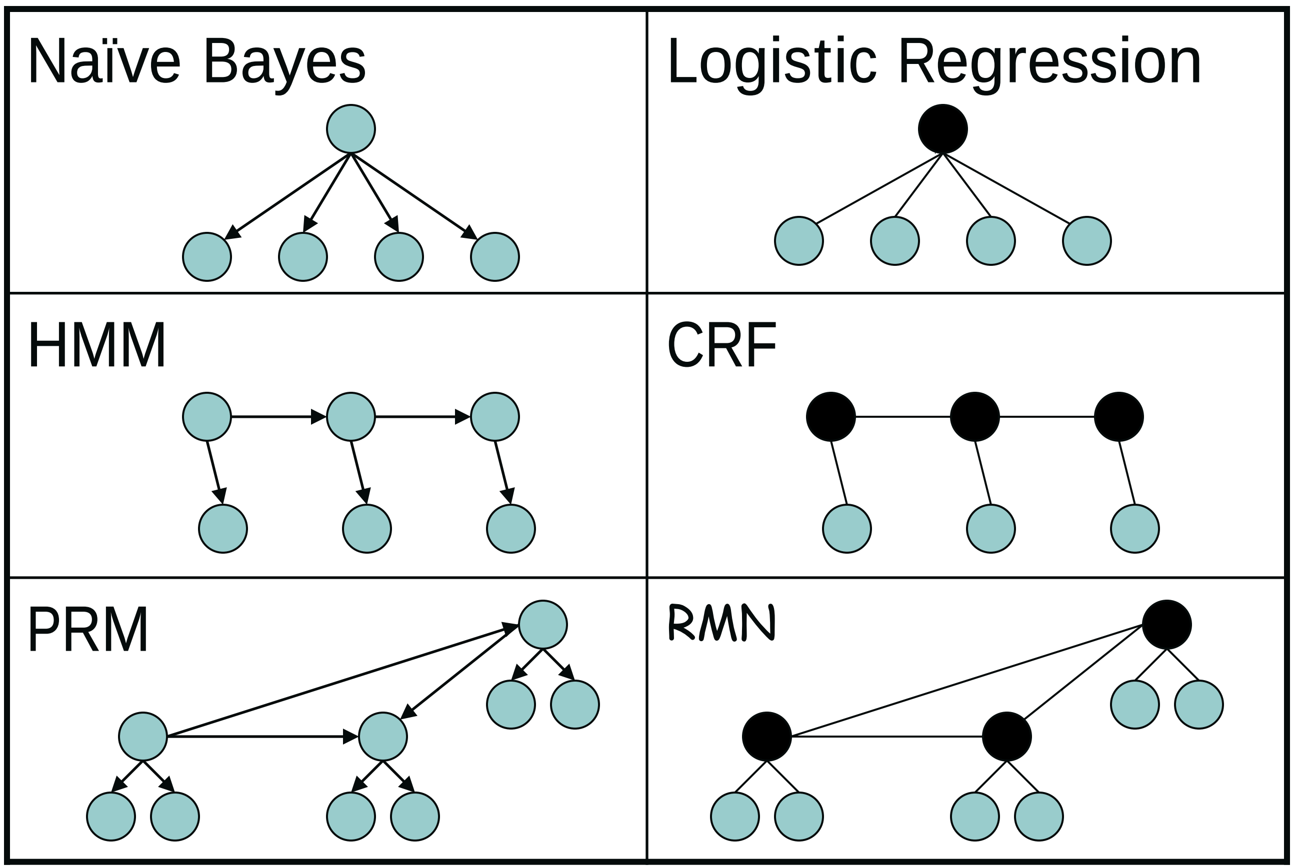
- A solution: model link structure
  - Exists uncertainty (Getoor+al. 2001)
  - Introduces v-structures for links



Problems with Directed Models

- Computational:
  - non-existent links also introduce a v-structure $O(n^2)$
- Representational:
  - acyclicity constraint does not allow many interesting relational dependencies
- Generative:
  - Learning objective not well-suited for classification

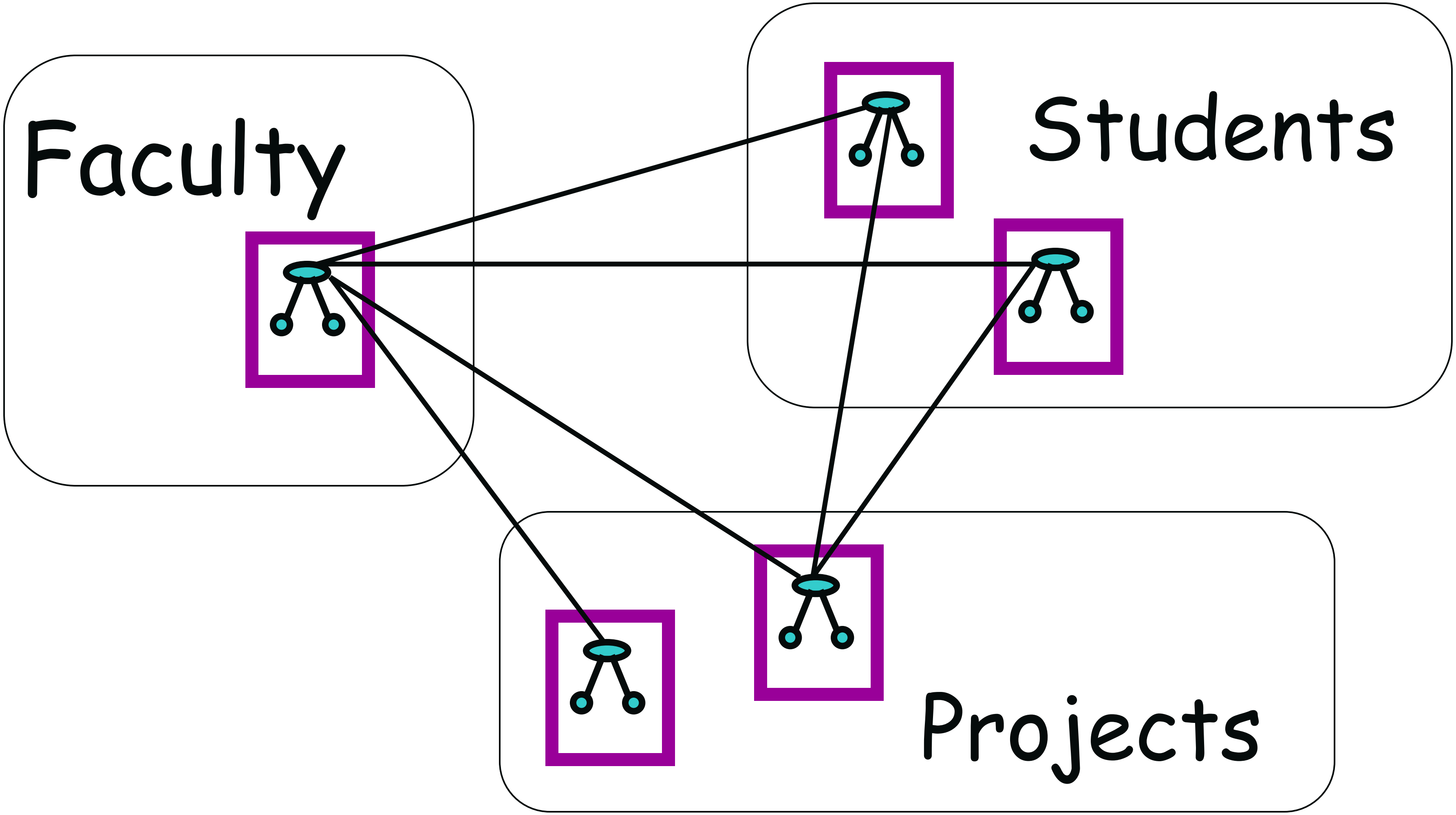
Directed vs Undirected



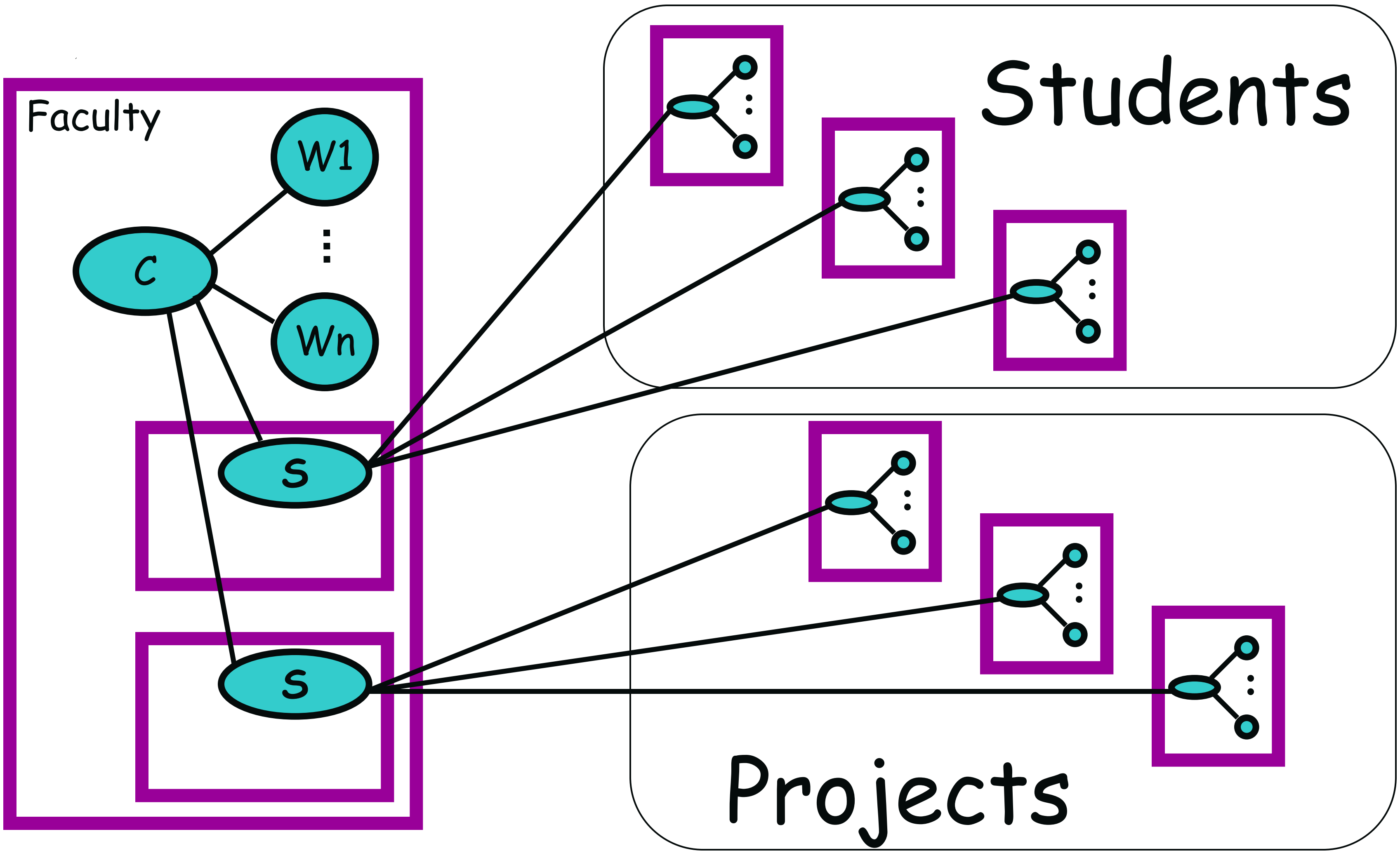
Markov networks

- Define distribution by a graph with potentials
  - $P(Y) = 1/Z \ \prod_c \phi_C(Y_C)$
- Conditional Markov networks
  - $P(Y|X) = 1/Z(X) \prod_c \phi_C(X_C,Y_C)$
- Log-linear representation of potentials
  - $\phi_C(X_C,Y_C) = \exp\{w_C \bullet f_C(X_C,Y_C)\}$
  - f_C – set of arbitrary features (e.g. indicator funcs)
  - $P(Y|X) = 1/Z(X) \exp\{w \bullet f(X,Y)\}$

Example RMN: Link model



Example RMN: Section model



Relational Domains

- Relational schema $\mathcal{E}$ defines a set of entity types:
  - $\mathcal{E} = E_1 \ldots E_n$
- Each entity type E specifies:
  - Content attributes E.X (Page.Has_word_blah)
  - Label attributes E.Y (Page.Category)
  - Reference attributes E.R (Link.From, Link.To)
- Instantiation I of a schema specifies:
  - a set of entities of each type
  - values of all attributes of all entities
  - $I = I(E_1) \ldots I(E_n)$

Relational Features

- Relational Neighborhood:
  - $R_I(e)$ = graph of entities reachable from e in I (subgraph of I rooted at e)
- Relational Feature:
  - $f(R_I(e)): R_I(e) \rightarrow \mathcal{R}$
  - Example:
    - link.from.category = link.to.category

RMNs

- RMN defines:
  - For each entity type E
    - a set of relational features f_E
    - a set of weights w_E
  - Conditional distribution for any instantiation I

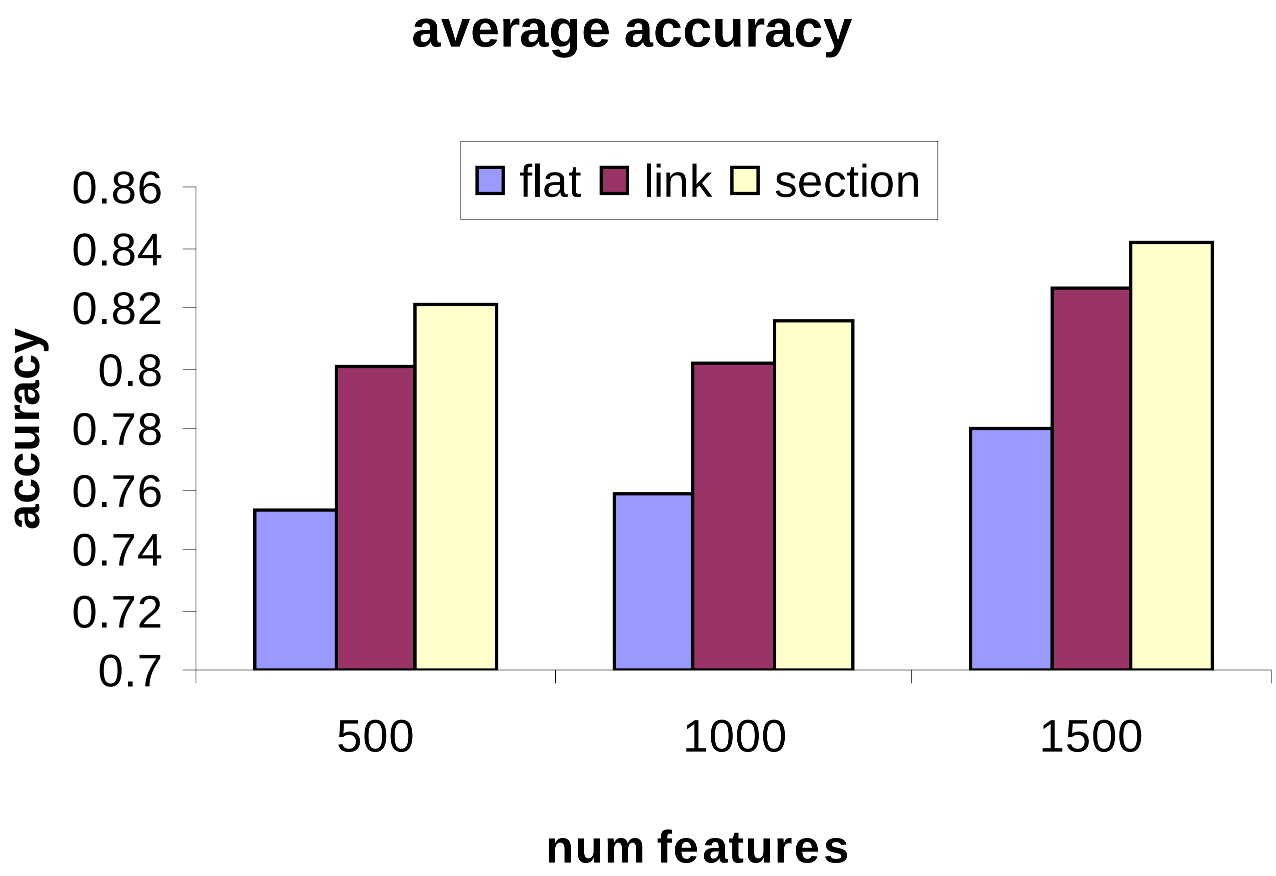
$$P(I.y|I.x,I.r) =$$

$$1/Z(I.x,I.r)\ \exp\{\ \Sigma_E\ \Sigma_{e \in I(E)}\ w_E \cdot f_E(R_I(e))\ \}$$

Learning

- Maximize conditional likelihood P(Y|X) + prior
- Convex optimization problem
- Conjugate gradient
  - Computing gradient requires inference
  - Exact inference intractable -> loopy belief propagation

Results: WebKB



Conclusion

Undirected models allow very flexible and rich models of relational data to be

- Represented compactly
- Learned from data discriminatively
- Used effectively for classification