Iterative Ensemble Classification for Relational Data

A Case Study of Semantic Web Services

Andreas Heß and Nicholas Kushmerick
University College Dublin, Ireland

September 18, 2004

Outline

1. Introduction
   - Relational Learning (simplified)
   - Motivation: Semantic Web Services
   - Relational Learning for Web Services
2. Iterative Ensemble Classification
   - Iterative Algorithms
   - Specialised Classifiers
   - Evaluation
3. Conclusion
   - Summary
   - Current and Future Work
   - Discussion

Outline

1. Introduction
   - Relational Learning (simplified)
   - Motivation: Semantic Web Services
   - Relational Learning for Web Services
2. Iterative Ensemble Classification
   - Iterative Algorithms
   - Specialised Classifiers
   - Evaluation
3. Conclusion
   - Summary
   - Current and Future Work
   - Discussion

Relational Learning

Relational Data

- consists of objects and relations between objects
- can be represented as a graphs

Three Types of Learning Tasks (following Slattery)

- classify nodes
- classify graphs
- classify subgraphs

Relational Learning

Relational Data

- consists of objects and relations between objects
- can be represented as a graphs

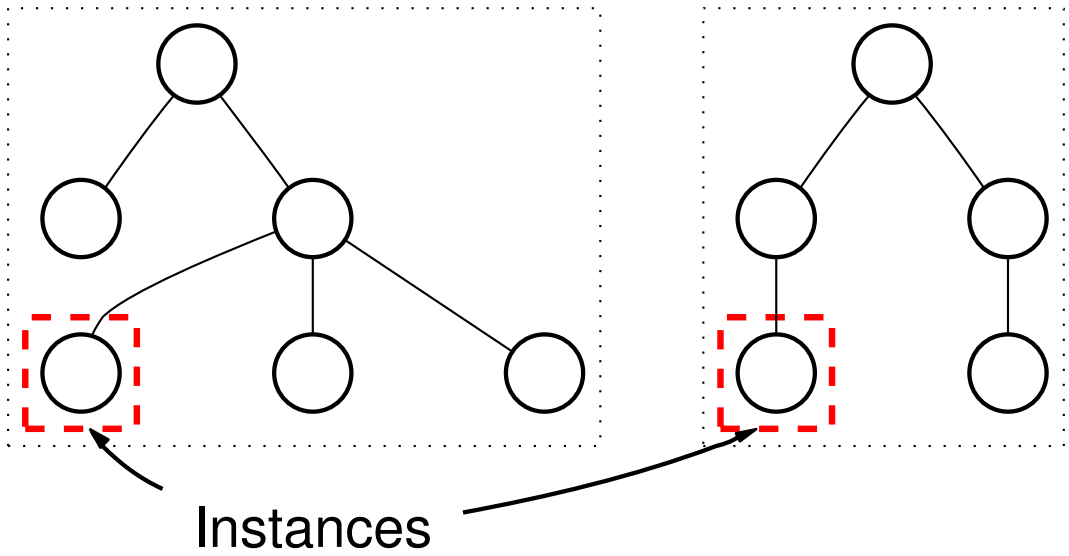
Three Types of Learning Tasks (following Slattery)

- classify nodes
- classify graphs
- classify subgraphs

Classifying Nodes

Task

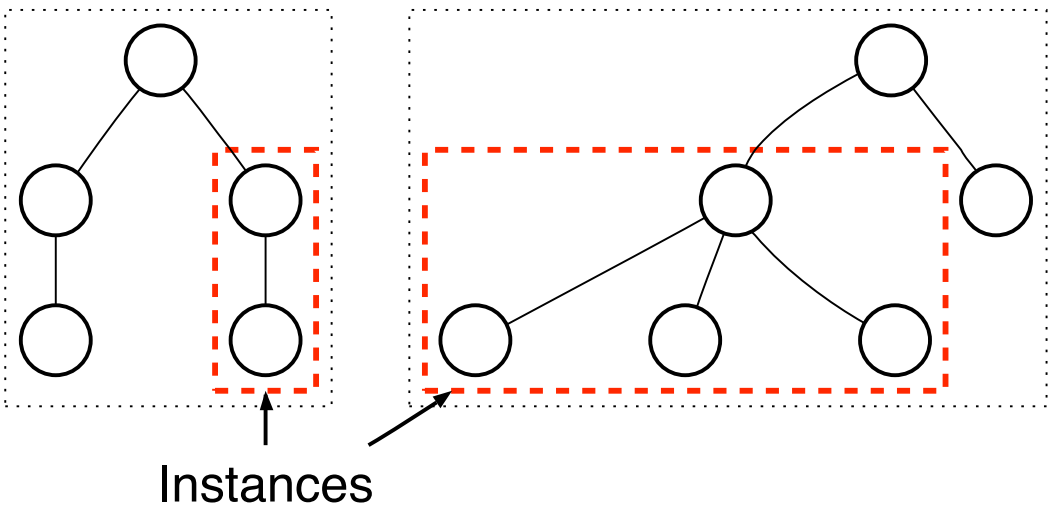
- Learn labels for nodes



Classifying Subgraphs

Task

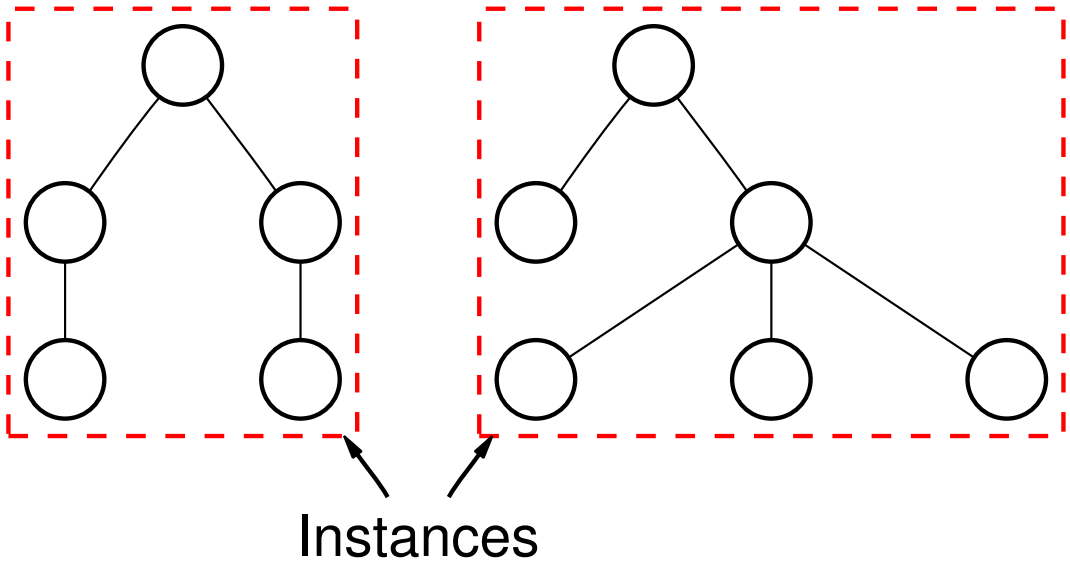
- Learn labels for subgraphs



Classifying Graphs

Task

- Learn labels for graphs



Relational Learning: Examples and Methods

Examples for Relational Learning

- Classical task: classify web pages

Methods for relational learning

- Iterative algorithms
- Statistical methods

Two Views

Intrinsic View

- Features inherent to instance
- e.g. text from web page

Extrinsic View

- Relations between instances
- e.g. class labels of linked web pages

(Following Neville and Jensen)

Two Views

Intrinsic View

- Features inherent to instance
- e.g. text from web page

Extrinsic View

- Relations between instances
- e.g. class labels of linked web pages

(Following Neville and Jensen)

Two Views

Intrinsic View

- Features inherent to instance
- e.g. text from web page

Extrinsic View

- Relations between instances
- e.g. class labels of linked web pages

(Following Neville and Jensen)

Two Views

Intrinsic View

- Features inherent to instance
- e.g. text from web page

Extrinsic View

- Relations between instances
- e.g. class labels of linked web pages

(Following Neville and Jensen)

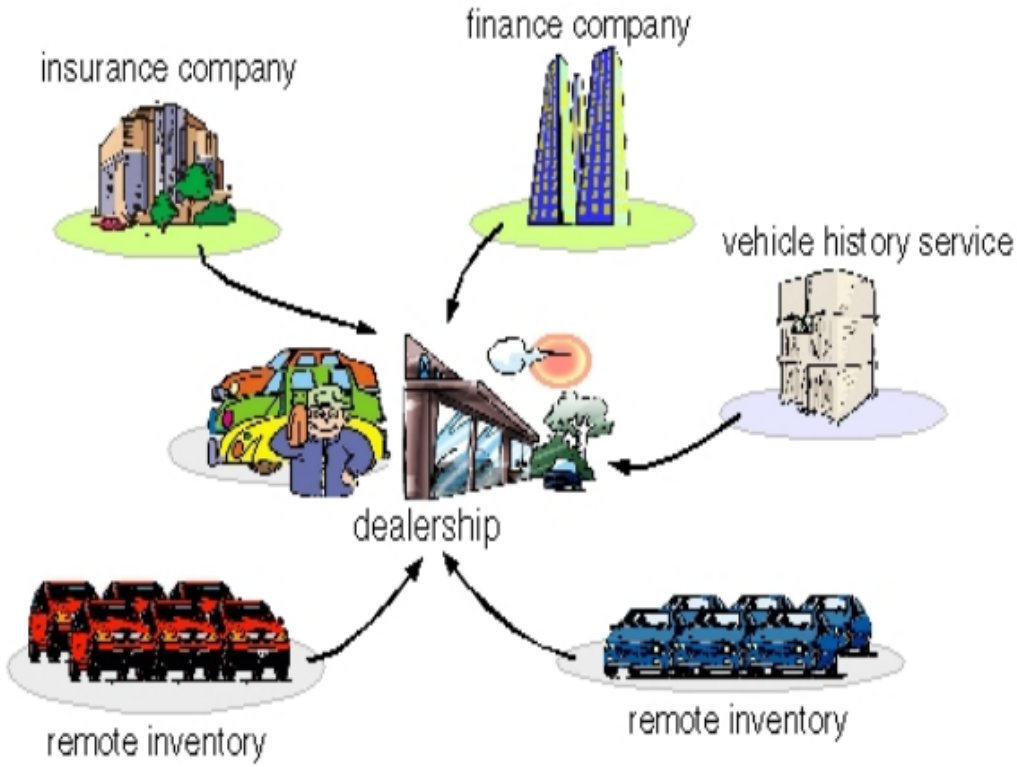
Now for Something Completely Different

1. Introduction
   - Relational Learning (simplified)
   - Motivation: Semantic Web Services
   - Relational Learning for Web Services
2. Iterative Ensemble Classification
   - Iterative Algorithms
   - Specialised Classifiers
   - Evaluation
3. Conclusion
   - Summary
   - Current and Future Work
   - Discussion

Web Services

Web Services

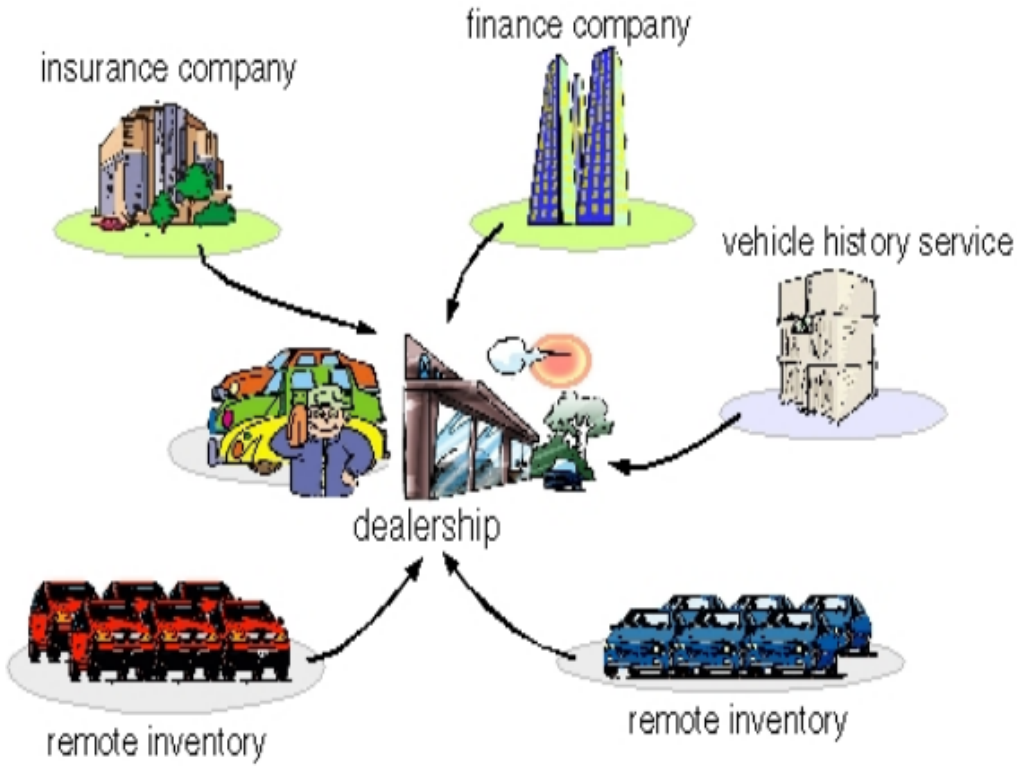
- Web-accessible software
- XML (SOAP) over HTTP
- Just RPC? Forms?
- Data Integration?



Web Service Descriptions (WSDL)

Web Services consist of:

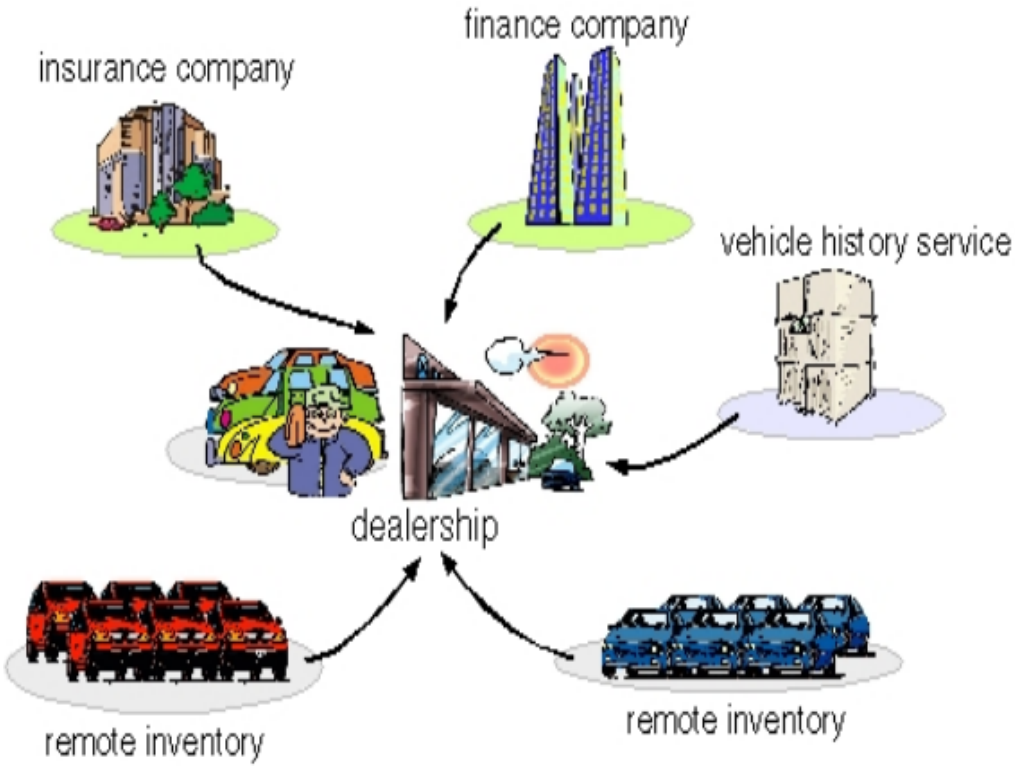
- Operations (methods)
- Messages (parameters)
- Complex types (structures)



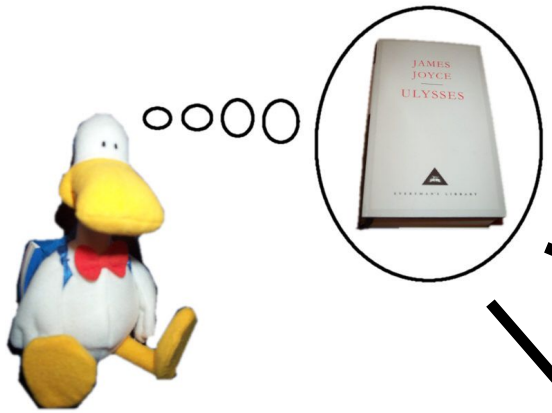
Semantic Web Services

Desired Features

- Automatic discovery
- Automatic composition
- Automatic invocation



Simple Scenario



Scenario:
Buying a book

?

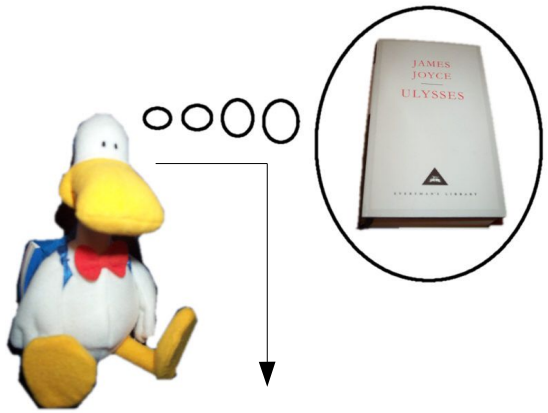
Congo
- author
- title
- quantity

?

WindingStair
- authName
- bookT
- ISBN

Teatime
- region
- qlty
- qty

Simple Scenario



Global Ontology

- Item
  - Quantity
  - Price
  - Book
    - Author
    - Title
    - ISBN
  - Tea
    - Region
    - Quality

Congo

- author
- title
- quantity

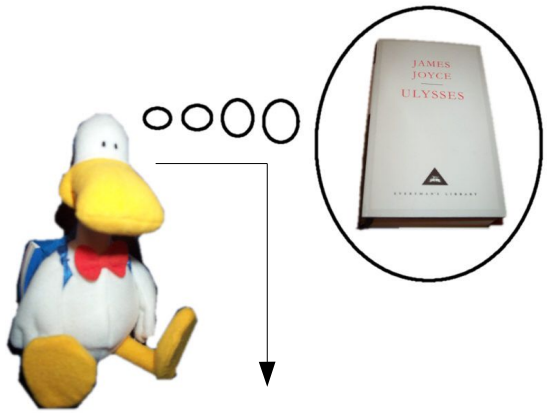
WindingStair

- authName
- bookT
- ISBN

Teatime

- region
- qlty
- qty

Simple Scenario



Semantic Metadata

(handcrafted)

Global Ontology

- Item
  - Quantity
  - Price
  - Book
    - Author
    - Title
    - ISBN
  - Tea
    - Region
    - Quality

Congo

- author
- title
- quantity

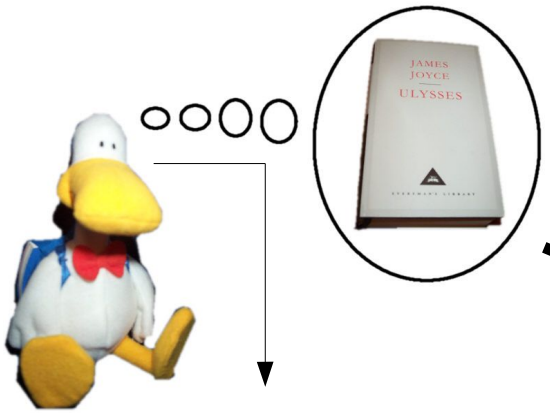
WindingStair

- authName
- bookT
- ISBN

Teatime

- region
- qlty
- qty

Simple Scenario



!

Congo

- author
- title
- quantity

WindingStair

- authName
- bookT
- ISBN

Teatime

- region
- qlty
- qty

Global Ontology

- Item
  - Quantity
  - Price
  - Book
    - Author
    - Title
    - ISBN
  - Tea
    - Region
    - Quality

Semantic Metadata

Assumptions

- Semantic annotation
- Shared ontology

Problem

- Hand-crafting annotations can be tedious

Our Solution

- Learn mappings from text to ontology

Semantic Metadata

Assumptions

- Semantic annotation
- Shared ontology

Problem

- Hand-crafting annotations can be tedious

Our Solution

- Learn mappings from text to ontology

Semantic Metadata

Assumptions

- Semantic annotation
- Shared ontology

Problem

- Hand-crafting annotations can be tedious

Our Solution

- Learn mappings from text to ontology

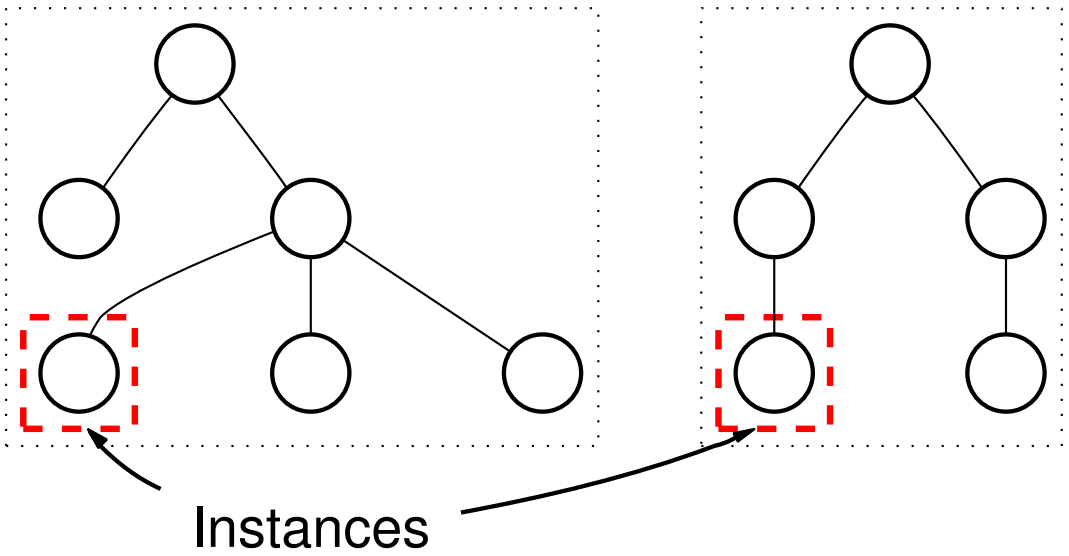
Outline

1. Introduction
   - Relational Learning (simplified)
   - Motivation: Semantic Web Services
   - Relational Learning for Web Services
2. Iterative Ensemble Classification
   - Iterative Algorithms
   - Specialised Classifiers
   - Evaluation
3. Conclusion
   - Summary
   - Current and Future Work
   - Discussion

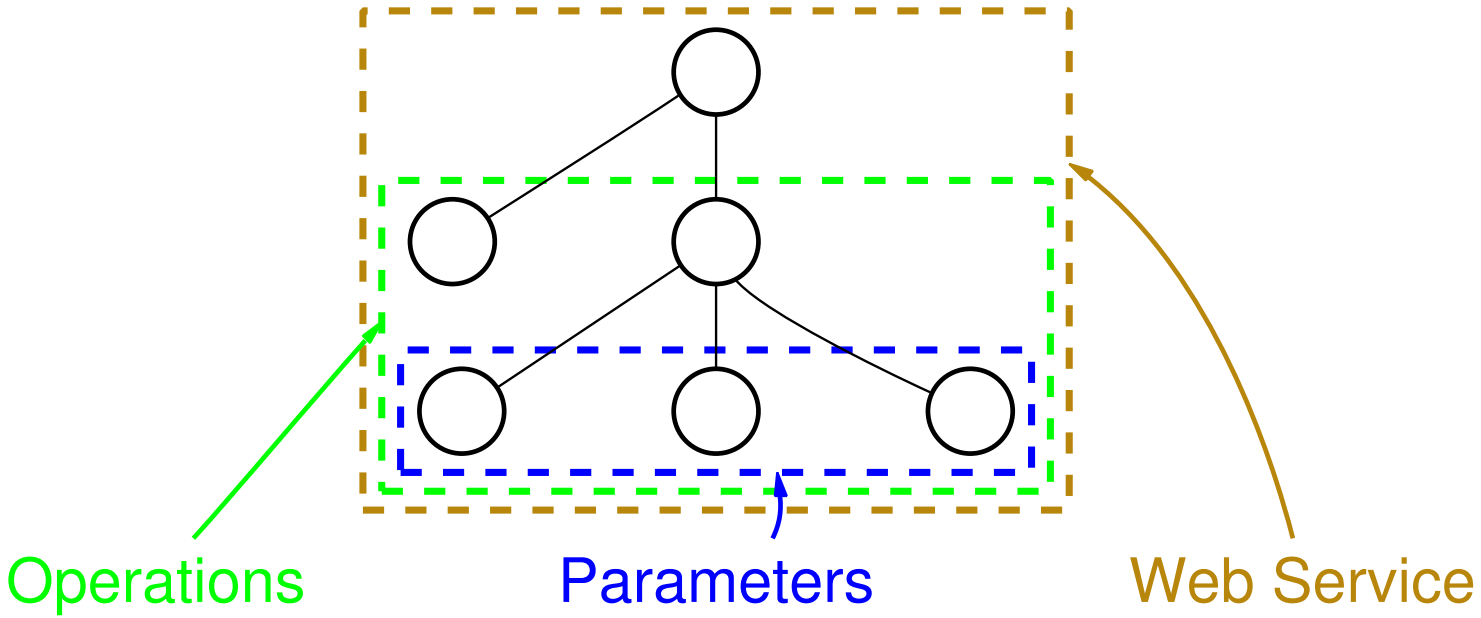
Recall: Classifying Nodes

Task

- Learn labels for nodes



Mapping Web Services to Relational Learning



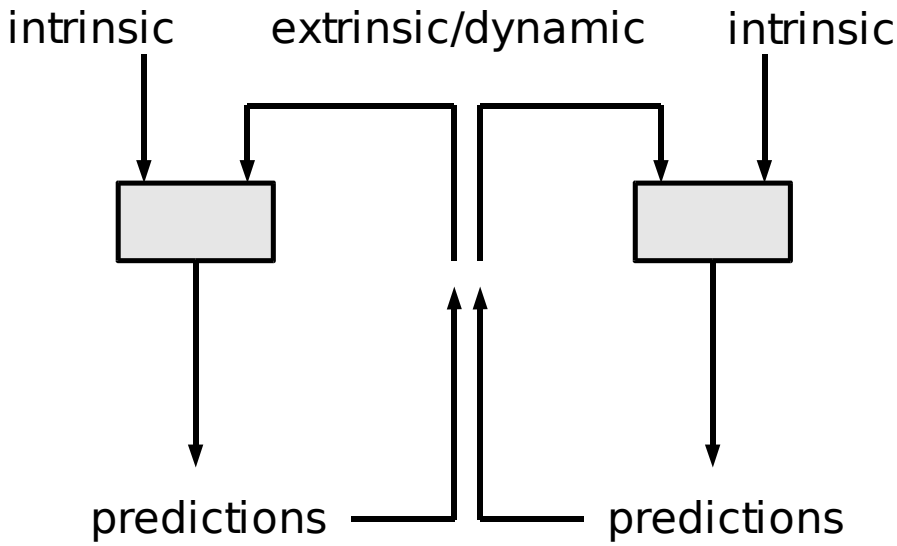
Outline

1. Introduction
   - Relational Learning (simplified)
   - Motivation: Semantic Web Services
   - Relational Learning for Web Services
2. Iterative Ensemble Classification
   - Iterative Algorithms
   - Specialised Classifiers
   - Evaluation
3. Conclusion
   - Summary
   - Current and Future Work
   - Discussion

Iterative Algorithm for Relational Classification

Iterative Algorithm

- Features are combination of intrinsic and extrinsic
- Extrinsic view changes with new results
- Predictions are fed back
- Extrinsic view is *dynamic*

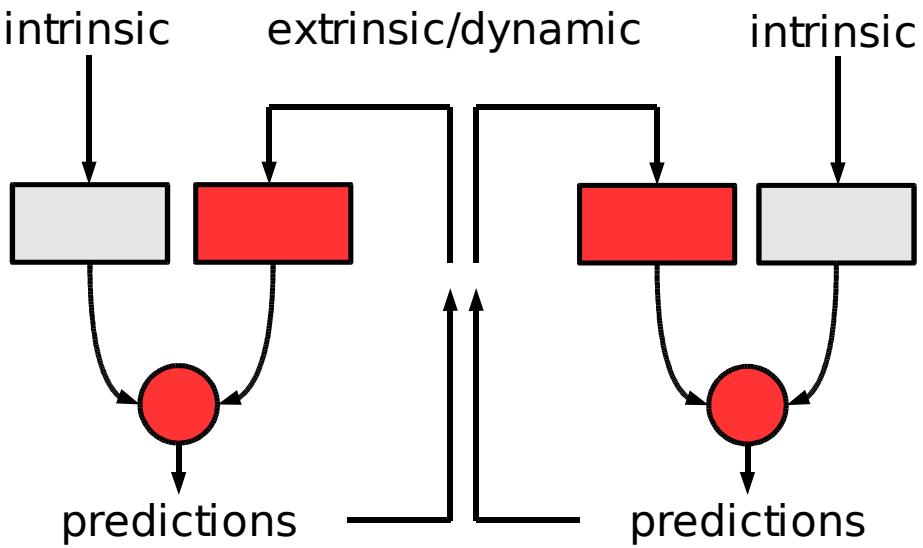


e.g. Neville/Jensen, Chakrabarti, Lu/Getoor

Iterative *Ensemble* Algorithm for Relational Classification

Iterative Ensemble Algorithm

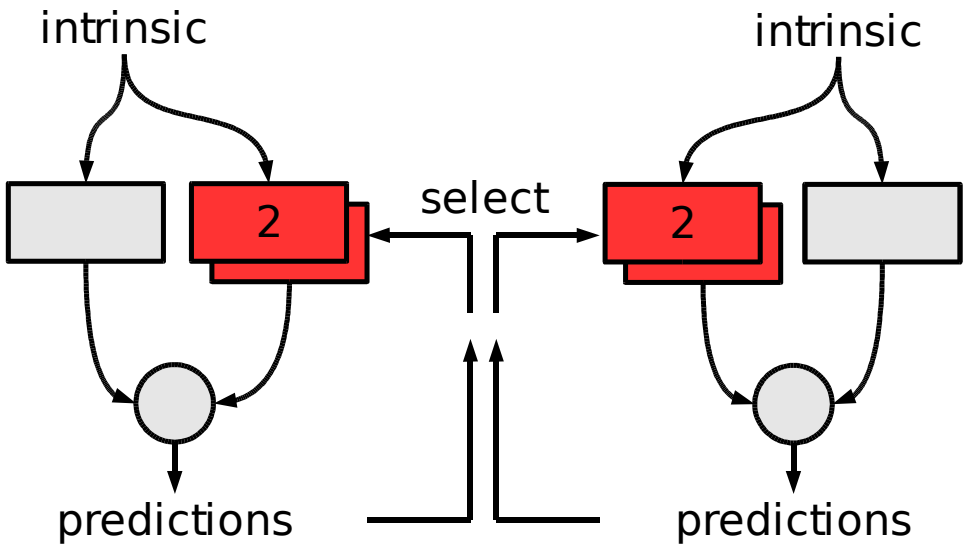
- Separate intrinsic/extrinsic classifiers
- Better performance with high-dimensional features



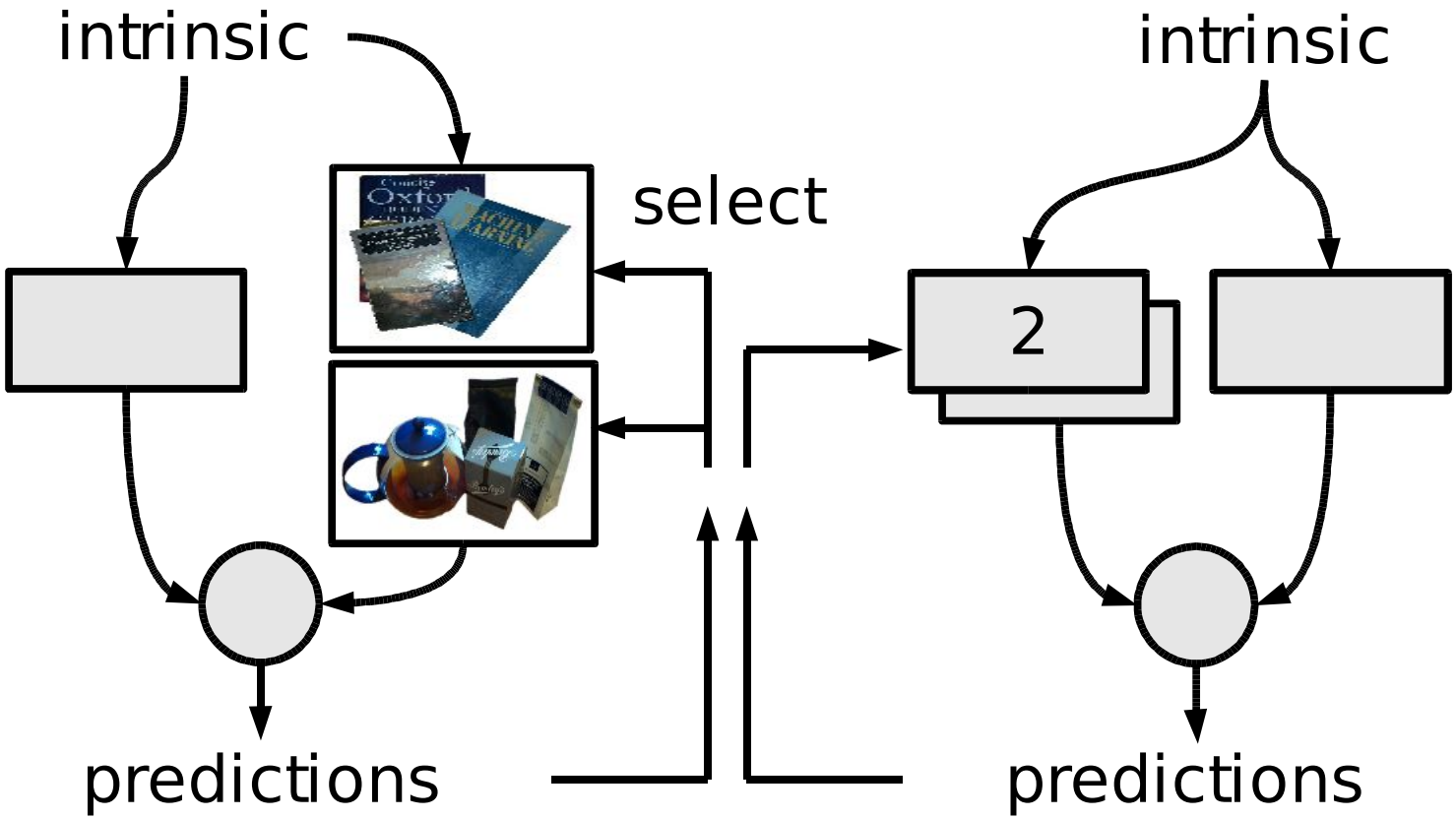
Iterative Ensemble Algorithm *with Specialised Classifiers*

Specialised Classifiers

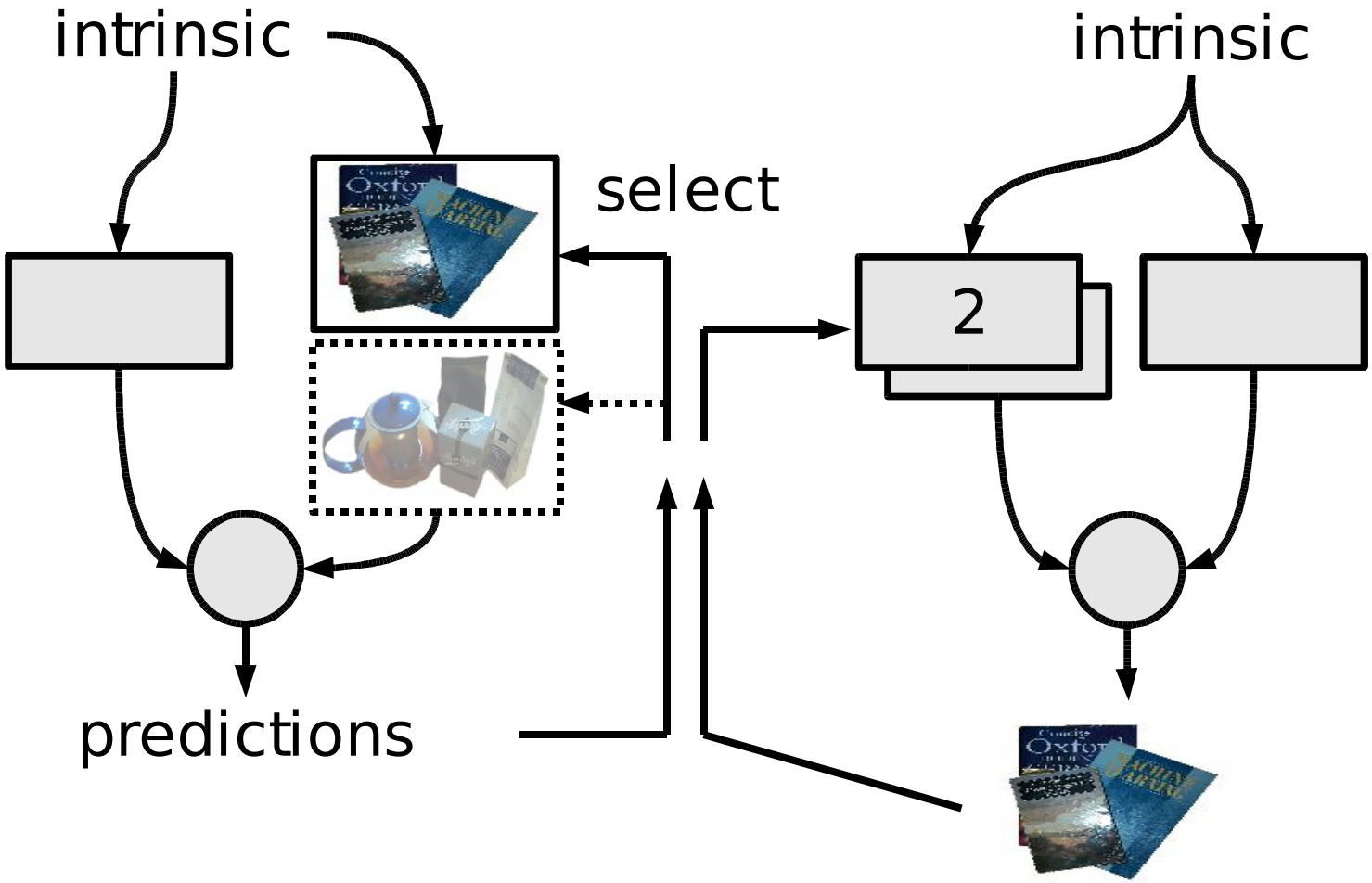
- Train several classifiers on intrinsic features
- Each one trained on subset of instances
- Classifier selected based on extrinsic features



Specialised Classifiers: Example



Specialised Classifiers: Example



Specialised Classifiers

Specialised Classifiers

- Useful if extrinsic view not sufficient on its own
- Extrinsic features serve as selector
- Idea can be applied in various ways
- e.g. Aidan Finn: Information Extraction (16:30h)

Specialised Classifiers

Specialised Classifiers

- Useful if extrinsic view not sufficient on its own
- Extrinsic features serve as selector
- Idea can be applied in various ways
- e.g. Aidan Finn: Information Extraction (16:30h)

Specialised Classifiers

Specialised Classifiers

- Useful if extrinsic view not sufficient on its own
- Extrinsic features serve as selector
- Idea can be applied in various ways
- e.g. Aidan Finn: Information Extraction (16:30h)

Specialised Classifiers

Specialised Classifiers

- Useful if extrinsic view not sufficient on its own
- Extrinsic features serve as selector
- Idea can be applied in various ways
- e.g. Aidan Finn: Information Extraction (16:30h)

Dataset

Our Web Services Dataset

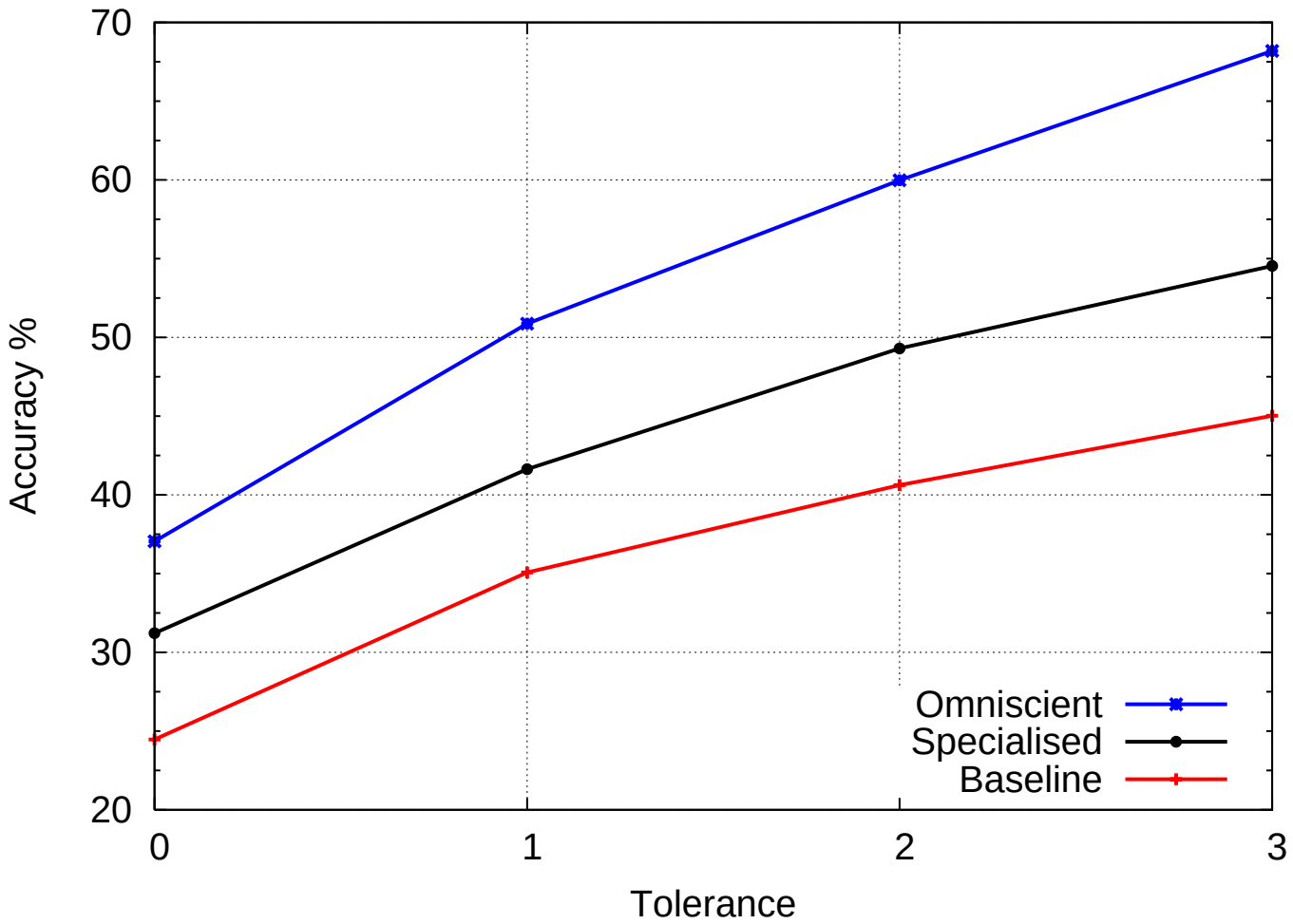
- 164 Web Services in 22 Categories
- 1138 Operations in 136 Domains
- 5452 Parameters with 312 Datatypes

Setup

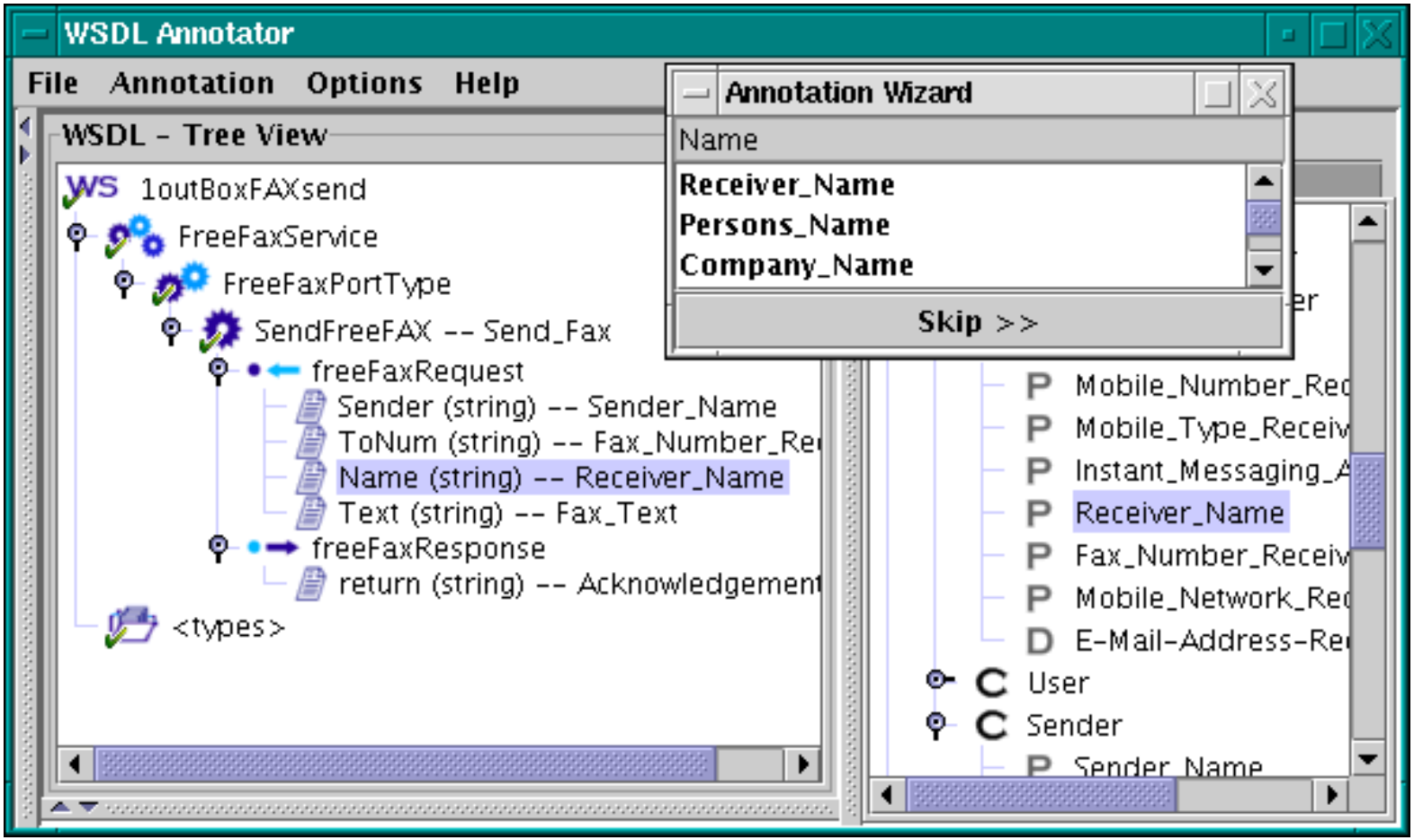
Experimental Setup

- Using iterative ensemble for service, operations
- Using specialised classifiers for datatypes
- Leave-one-service-out
- Compared to omniscient setup
  with extrinsic view always correct
- Compared to non-relational baseline
- Non-ensemble setting ommitted: always worse than baseline

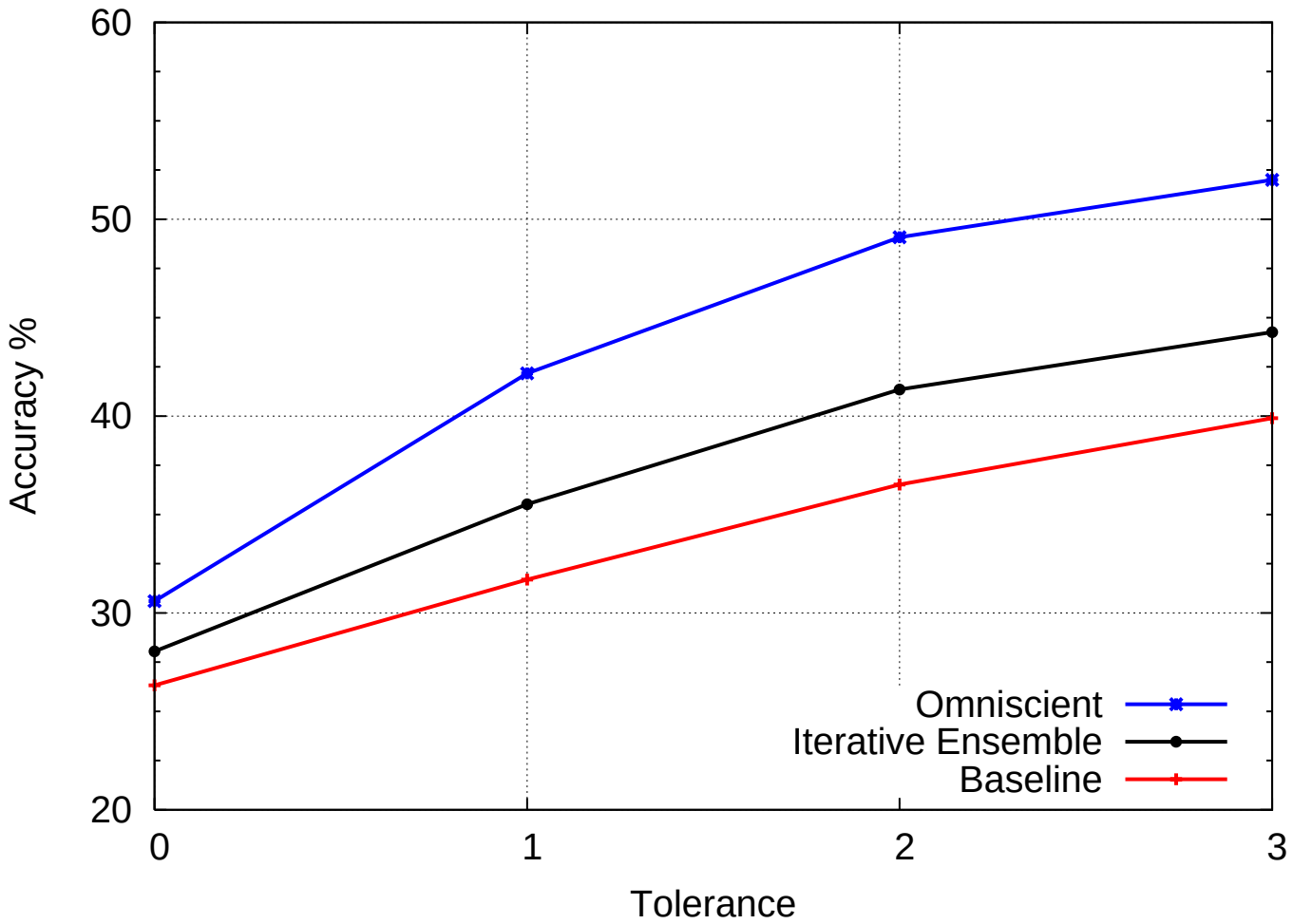
Evaluation: Datatype of Parameters



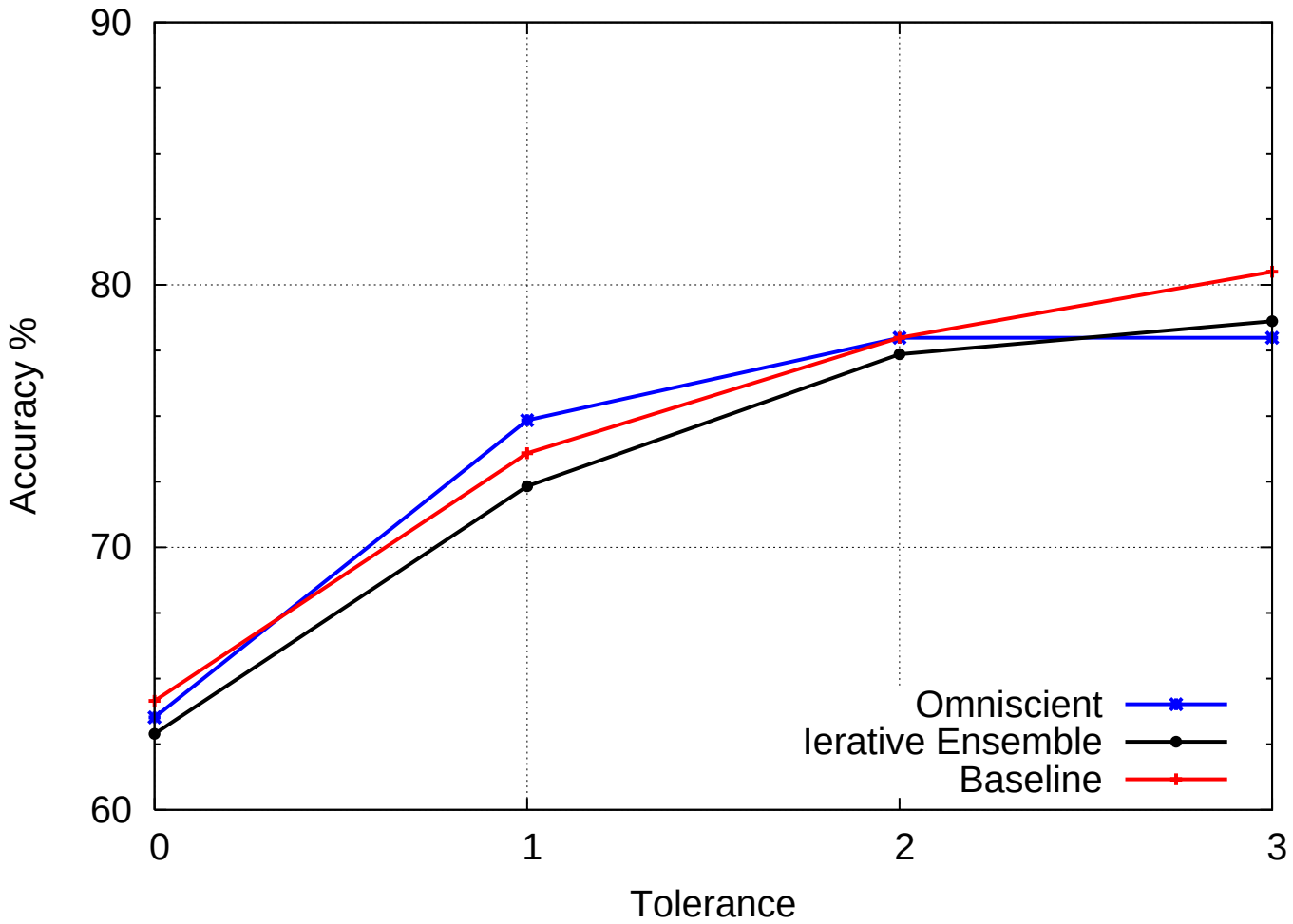
Why Allow Tolerance?



Evaluation: Domain of Operations



Evaluation: Category of Service



Outline

1. Introduction
   - Relational Learning (simplified)
   - Motivation: Semantic Web Services
   - Relational Learning for Web Services
2. Iterative Ensemble Classification
   - Iterative Algorithms
   - Specialised Classifiers
   - Evaluation
3. Conclusion
   - Summary
   - Current and Future Work
   - Discussion

Summary

We have presented...

- An iterative ensemble method for relational learning
- Specialised classifiers for relational learning

Evaluation has shown...

- If features high-dimensional:
  ensemble better than single classifier
- If extrinsic view alone not sufficient: specialised classifiers

Summary

We have presented...

- An iterative ensemble method for relational learning
- Specialised classifiers for relational learning

Evaluation has shown...

- If features high-dimensional:
  ensemble better than single classifier
- If extrinsic view alone not sufficient: specialised classifiers

Current and Future Work

Current and Future Work

- Post-processing result based on mined rules
- Improve the annotator application
- Eddie Johnston: Aggregating Web Service output
- See also our ISWC paper

Current and Future Work

Current and Future Work

- Post-processing result based on mined rules
- Improve the annotator application
- Eddie Johnston: Aggregating Web Service output
- See also our ISWC paper

Current and Future Work

Current and Future Work

- Post-processing result based on mined rules
- Improve the annotator application
- Eddie Johnston: Aggregating Web Service output
- See also our ISWC paper

Current and Future Work

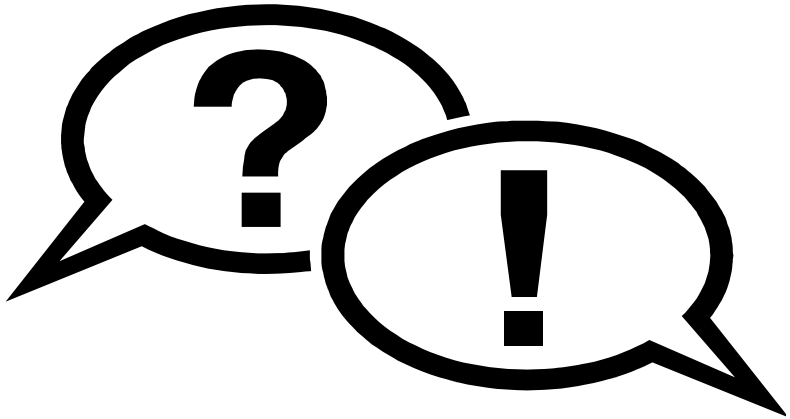
Current and Future Work

- Post-processing result based on mined rules
- Improve the annotator application
- Eddie Johnston: Aggregating Web Service output
- See also our ISWC paper

Discussion

Thank You for Your Attention

- Questions?



URLs and email addresses

- {andreas.hess,nick}@ucd.ie
- http://moguntia.ucd.ie, http://www.cs.ucd.ie/staff/nick/
- http://smi.ucd.ie/RSWS