# Heterogeneous Transfer Learning
## adapted from ACL'09 Invited Talk

Qiang Yang

Hong Kong University of Science and Technology
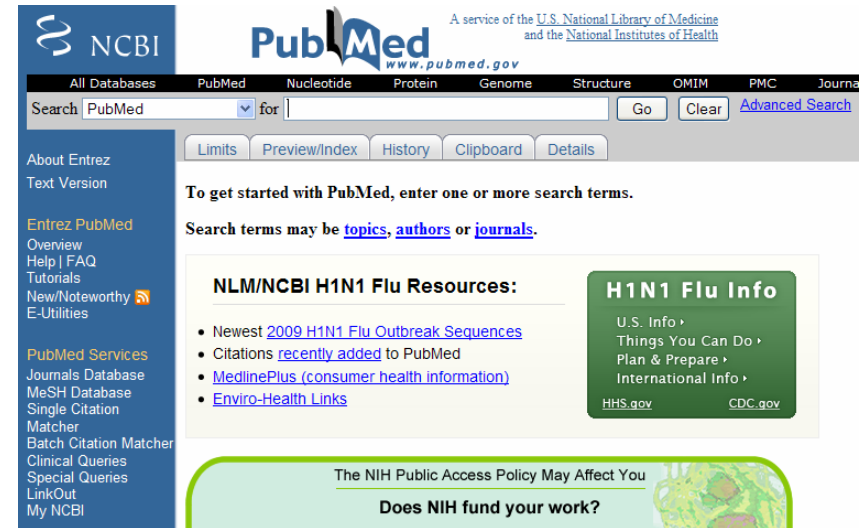Hong Kong, China

http://www.cse.ust.hk/~qyang

# A Major Assumption w/ Machine Learning

- Training and future (test) data
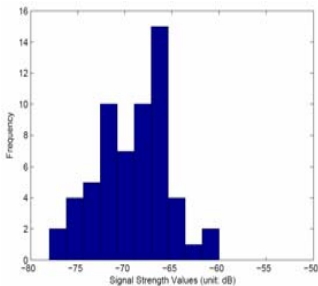  - follow the same distribution, and
  - are in same feature space

# When distributions are different



- Part-of-Speech tagging
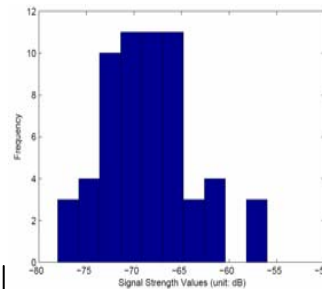- Named-Entity Recognition
- Classification

# When distributions are different

- Wireless sensor networks
- Different time periods, devices or space



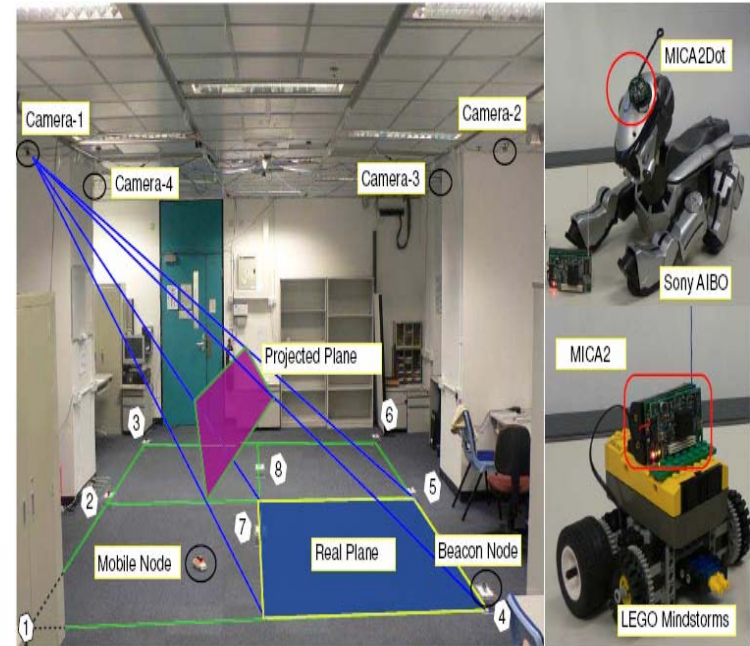- <span style="color:red">Relabeling data can be expensive</span>



Night time

Device 1

Day time

Device 2

Device, Space, or Time

# When Features are different

- Heterogeneous: different feature spaces

**Training: Text**  **Future: Images**

**Apples**

The apple is the pomaceous fruit of the apple tree, species Malus domestica in the rose family Rosaceae ...

**Bananas**

Banana is the common name for a type of fruit and also the herbaceous plants of the genus Musa which produce this commonly eaten fruit ...
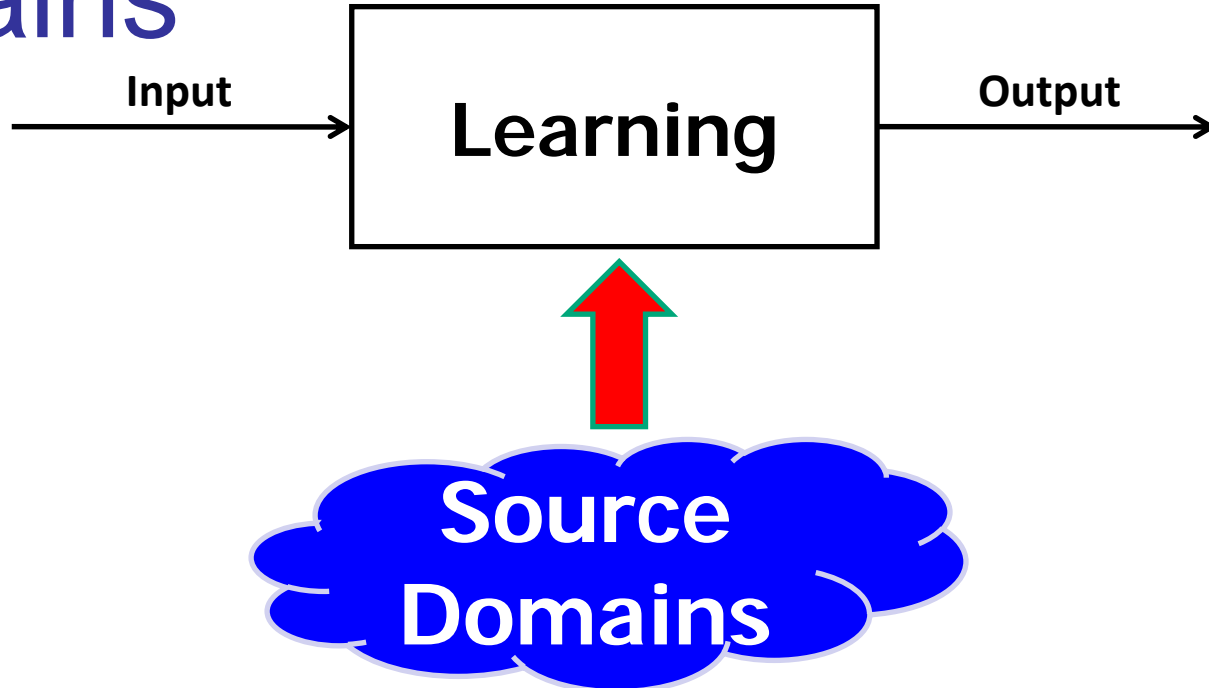
# Transfer Learning?

- People often transfer knowledge to novel situations
    - Chess → Checkers
    - C++ → Java
    - Physics → Computer Science

Transfer Learning:

The ability of a system to recognize and apply knowledge and skills learned in previous tasks to novel tasks (or new domains)

# Transfer Learning: Source Domains

Input → | Learning | → Output

**Source Domains**

|  | Source Domain | Target Domain |
|---|---|---|
| Training Data | Labeled/Unlabeled | Labeled/Unlabeled |
| Test Data |  | Unlabeled |

# Outline

- **Transfer Learning Basics**
- Homogeneous Transfer Learning
- Heterogeneous Transfer Learning
- Future Works

# Outline

- **Transfer Learning Basics**
- Homogeneous Transfer Learning
- Heterogeneous Transfer Learning
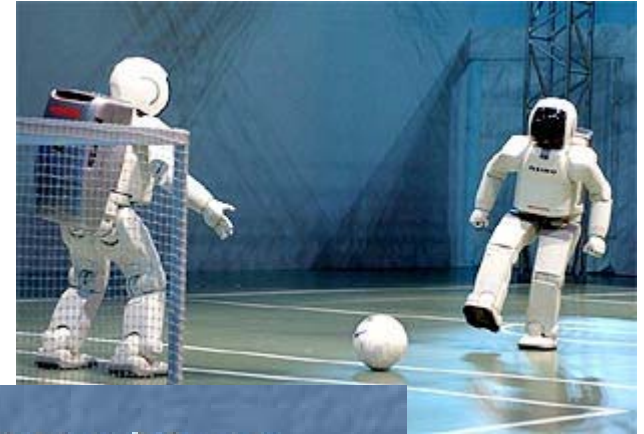- Future Works

**Transfer Learning Survey**
S. Pan and Q. Yang, *A Survey on Transfer Learning* *IEEE TKDE 2009.*
*http://www.cse.ust.hk/~sinnopan/SurveyTL.htm*

# Reinforcement Learning

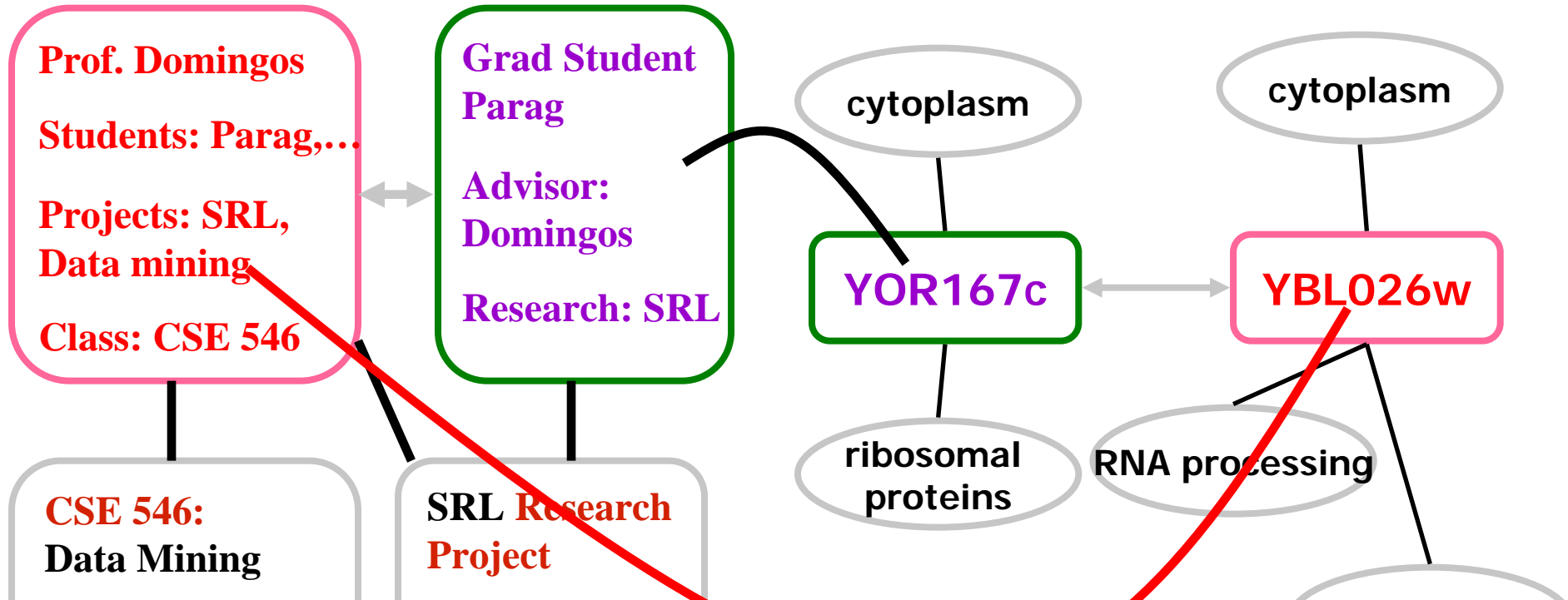**L. Torrey, J. Shavlik, S. Natarajan, P. Kuppili & T. Walker (2008).**
**Transfer in Reinforcement Learning via Markov Logic Networks**. *AAAI'08 Workshop on Transfer Learning for Complex Tasks*, Chicago, IL.

# Deep Transfer w/ Markov Logic Network [Davis and Domingos, ICML 2009]
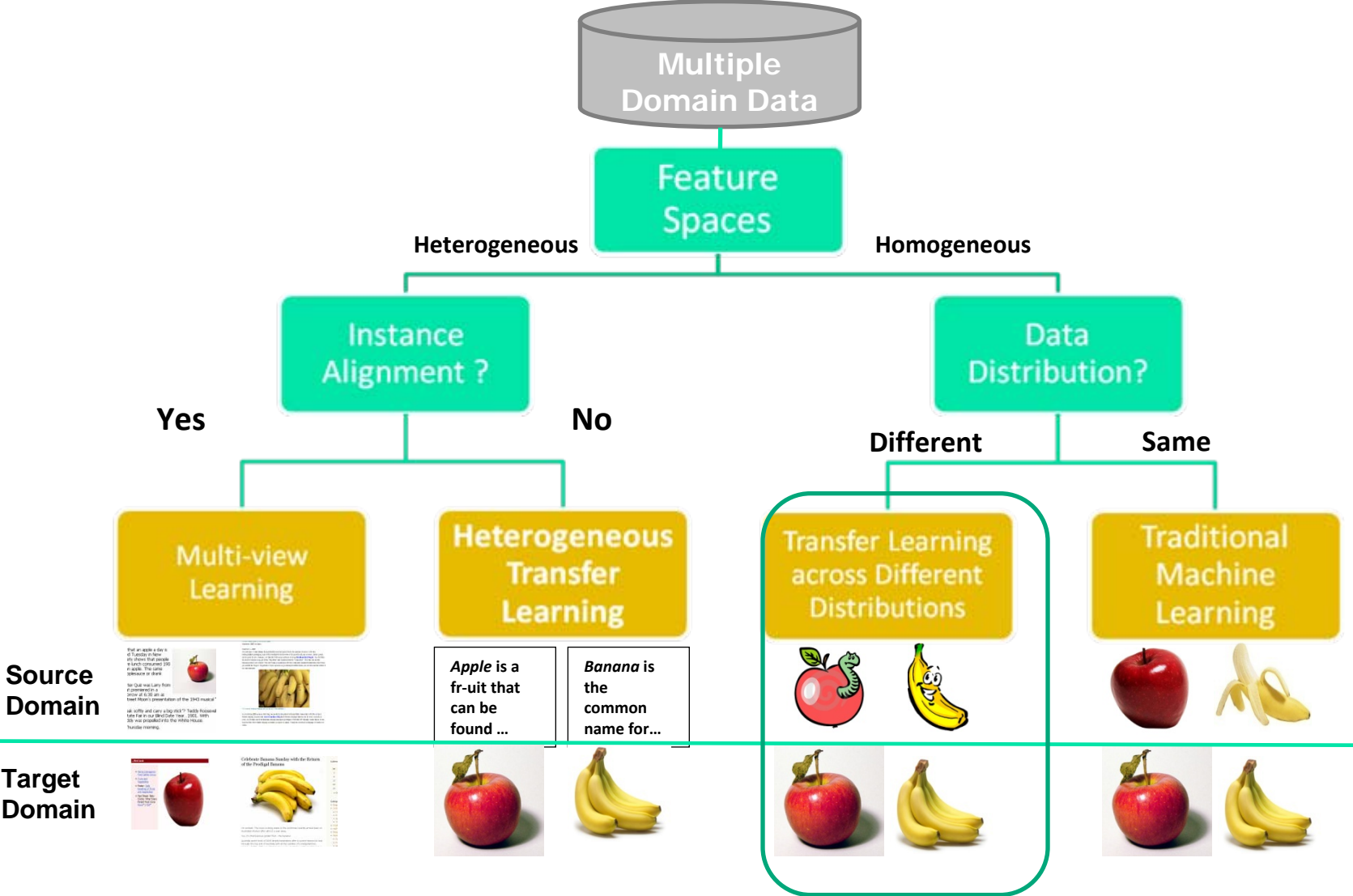
**Source Domain**

**Target Domain**

**Prof. Domingos**

**Students: Parag,...**

**Projects: SRL, Data mining**

**Class: CSE 546**

**Grad Student Parag**

**Advisor: Domingos**

**Research: SRL**

cytoplasm

cytoplasm

**YOR167c**

**YBL026w**

ribosomal proteins

RNA processing

**CSE 546:**
**Data Mining**

**SRL Research Project**

To

Ho

**Complex(z, y) ∧ Interacts(x, z)⇒ Complex(x, y) and**

**Location(z, y) ∧ Interacts(x, z) ⇒ Location(x, y)**

# Different Learning Problems



**Domain Adaptation**

# Domain Adaptation in NLP

**Applications**

- Automatic Content Extraction
- Sentiment Classification
- Part-Of-Speech Tagging
- NER
- Question Answering
- Classification
- Clustering

**Selected Methods**

- Domain adaptation for statistical classifiers [Hal Daume III & Daniel Marcu, JAIR 2006], [Jiang and Zhai, ACL 2007]
- Structural Correspondence Learning [John Blitzer et al. ACL 2007] [Ando and Zhang, JMLR 2005]
- Latent subspace [Sinno Jialin Pan et al. AAAI 08]
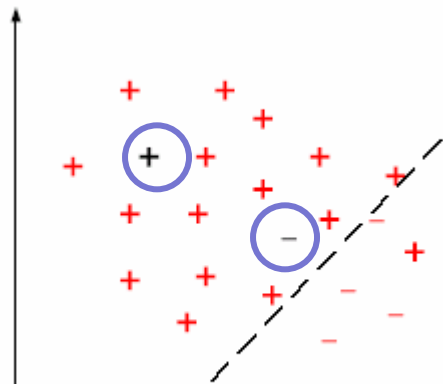
# Instance-transfer Ap

[Wu and Dietterich ICML-04]
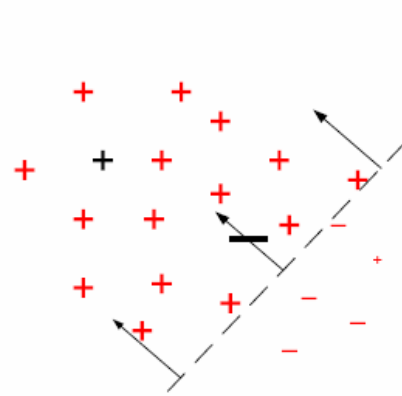[J.Jiang and C. Zhai, ACL 2007]
[Dai, Yang et al. ICML-07]

- Cross-domain POS tagging,
- entity type classification
- Personalized spam filtering

**Uniform weights**

**Correct the decision boundary by re-weighting**



**Loss function on the target domain data**

**Loss function on the source domain data**

**Regularization term**

➢ Differentiate the cost for misclassification of the target and source data

$$J(h) = \sum_{i}^{n_T} L(h(x_{T_i}), y_{T_i}) + \lambda \sum_{j}^{n_S} L(h(x_{S_j}), y_{T_j}) + R(h)$$

# TrAdaBoost

## [Dai, Yang et al. ICML-07]

**Here** [Fr

To decr
of the misclassified data

of the misclassified data

**Evaluation with 20NG: 22%→8%**
http://people.csail.mit.edu/jrennie/20Newsgroups/

- Misclassified
  - of the misclassified target data
  - decrease the weights of the misclassified source data

**Source domain labeled data**

**target domain labeled data**

**The whole training data set**

**Classifiers trained on re-weighted labeled data**
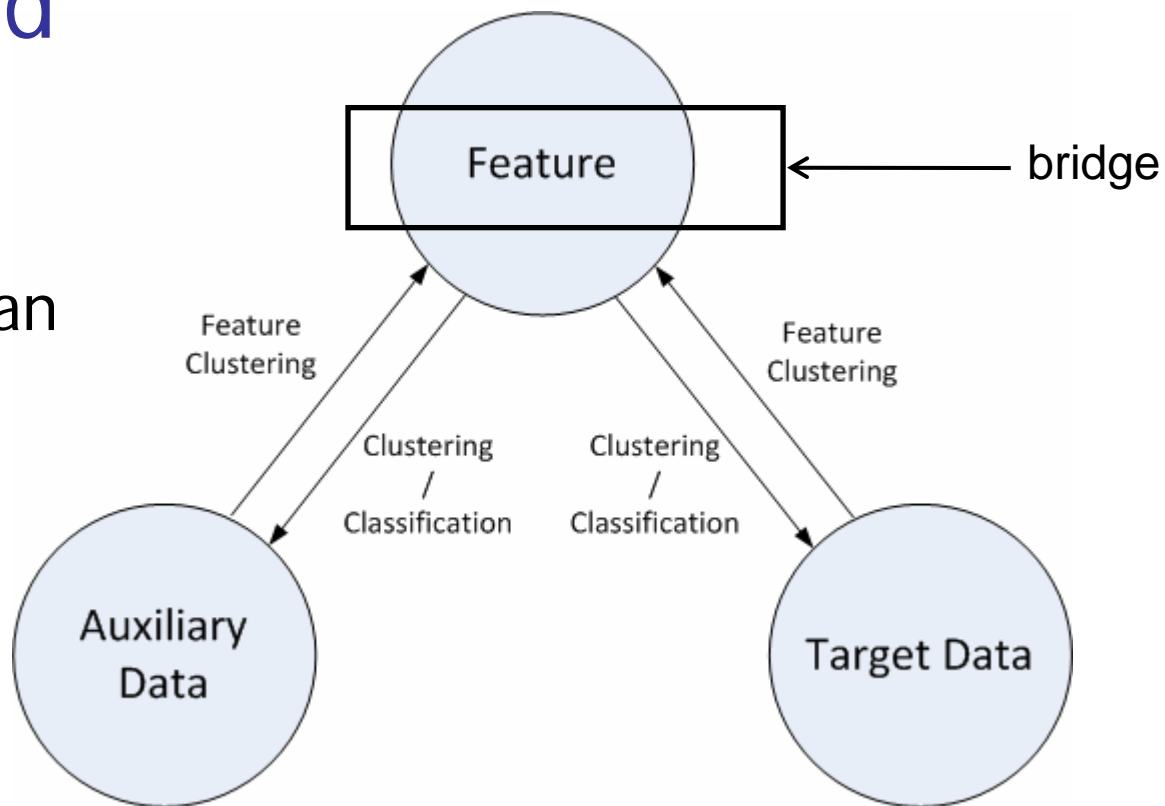
**Target domain unlabeled data**
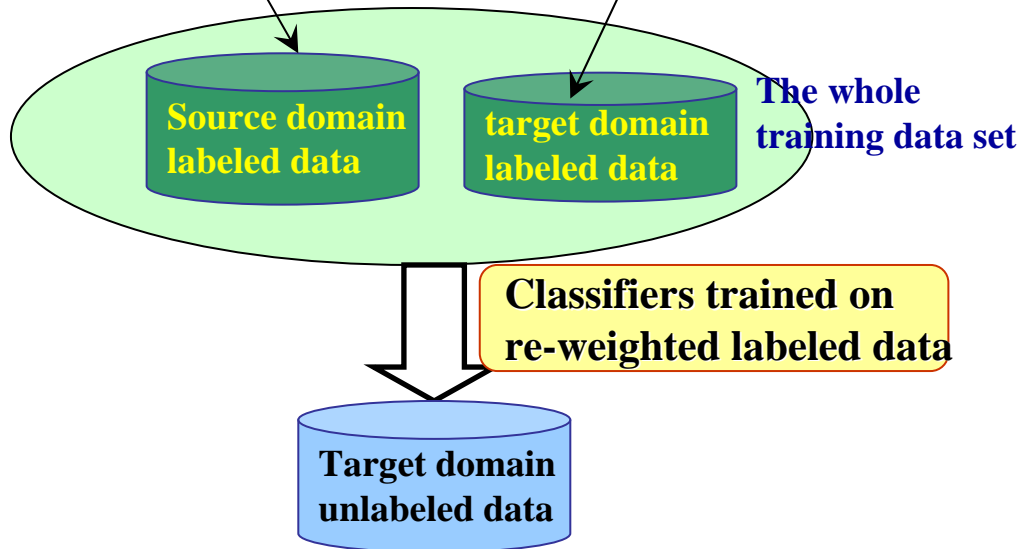
# Feature-based Transfer Learning
## [Dai, Xue, Yang et al. KDD 2007]

- # Target:
  - ## All unlabeled instances
- Distributions
  - Feature spaces can be different, but have overlap
  - Same classes
  - P(X,Y): different!

CoCC=Co-clustering based Classification

# Document-word co-occurrence

**Source**   D_S

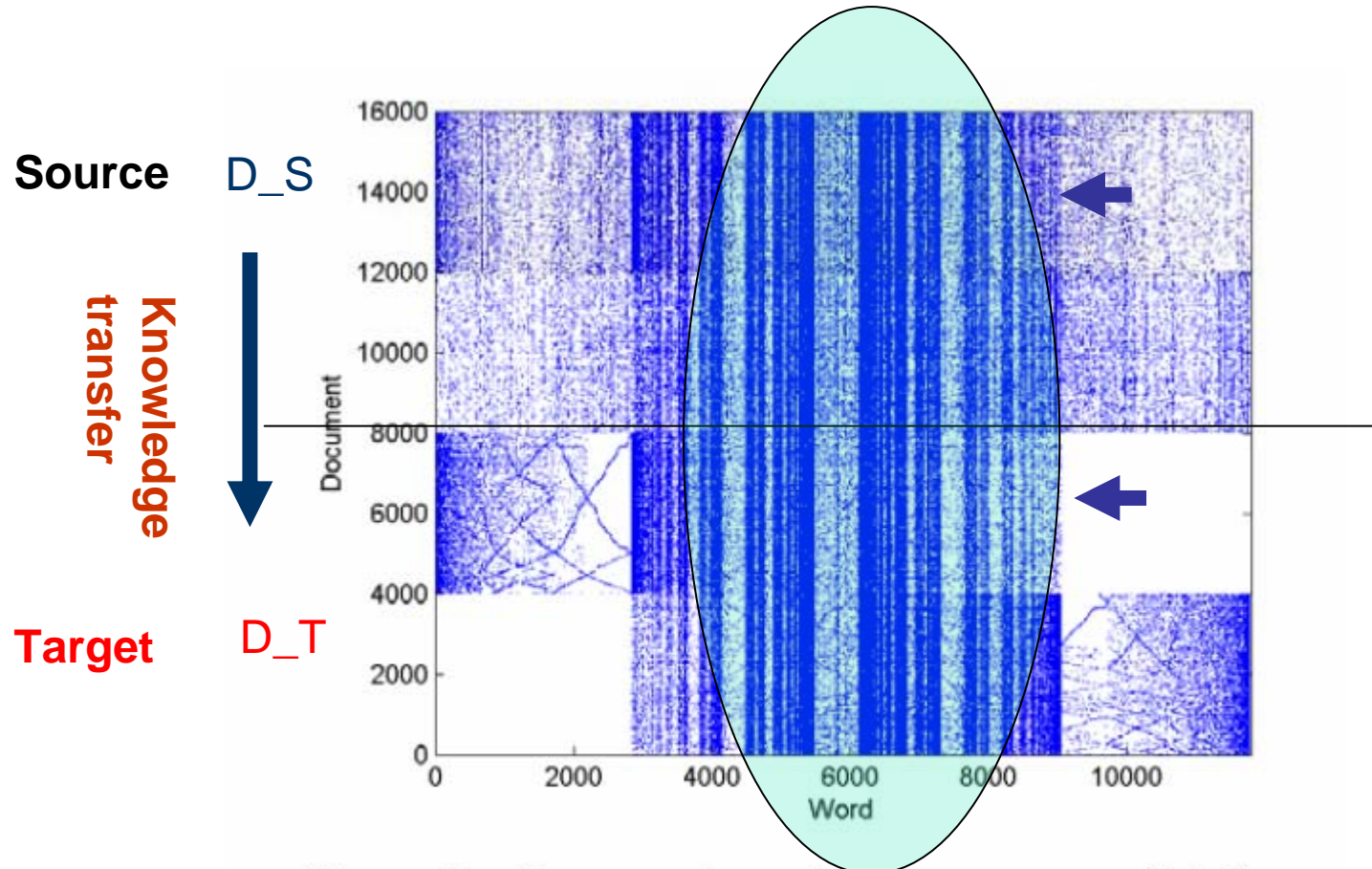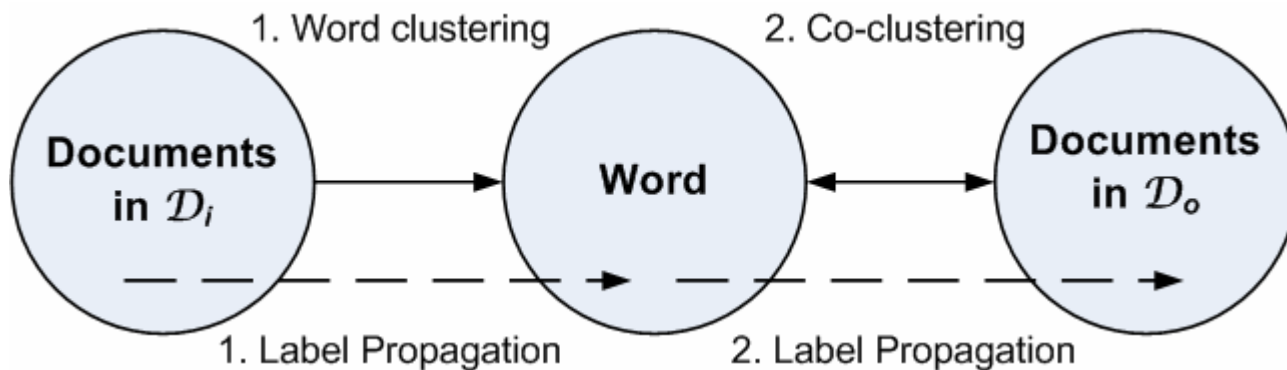**Knowledge transfer**

**Target**   D_T



Figure 2: Document-word co-occurrence distribution on the auto vs aviation data set

# Co-Clustering based transfer [Dai, Xue, Yang et al. KDD 2007]

- **Co-clustering is applied between**
  - features (words) and target-domain documents
  - constrained by the labels of source domain documents
  - word clusters in both domains: a ***bridge***

# Structural Correspondence Learning [Blitzer et al. ACL 2007]

- SCL: [Ando and Zhang, JMLR 2005]
- Method
  - Define pivot features: common in two domains
  - Find non-pivot features in each domain
  - Build classifiers through the non-pivot Features

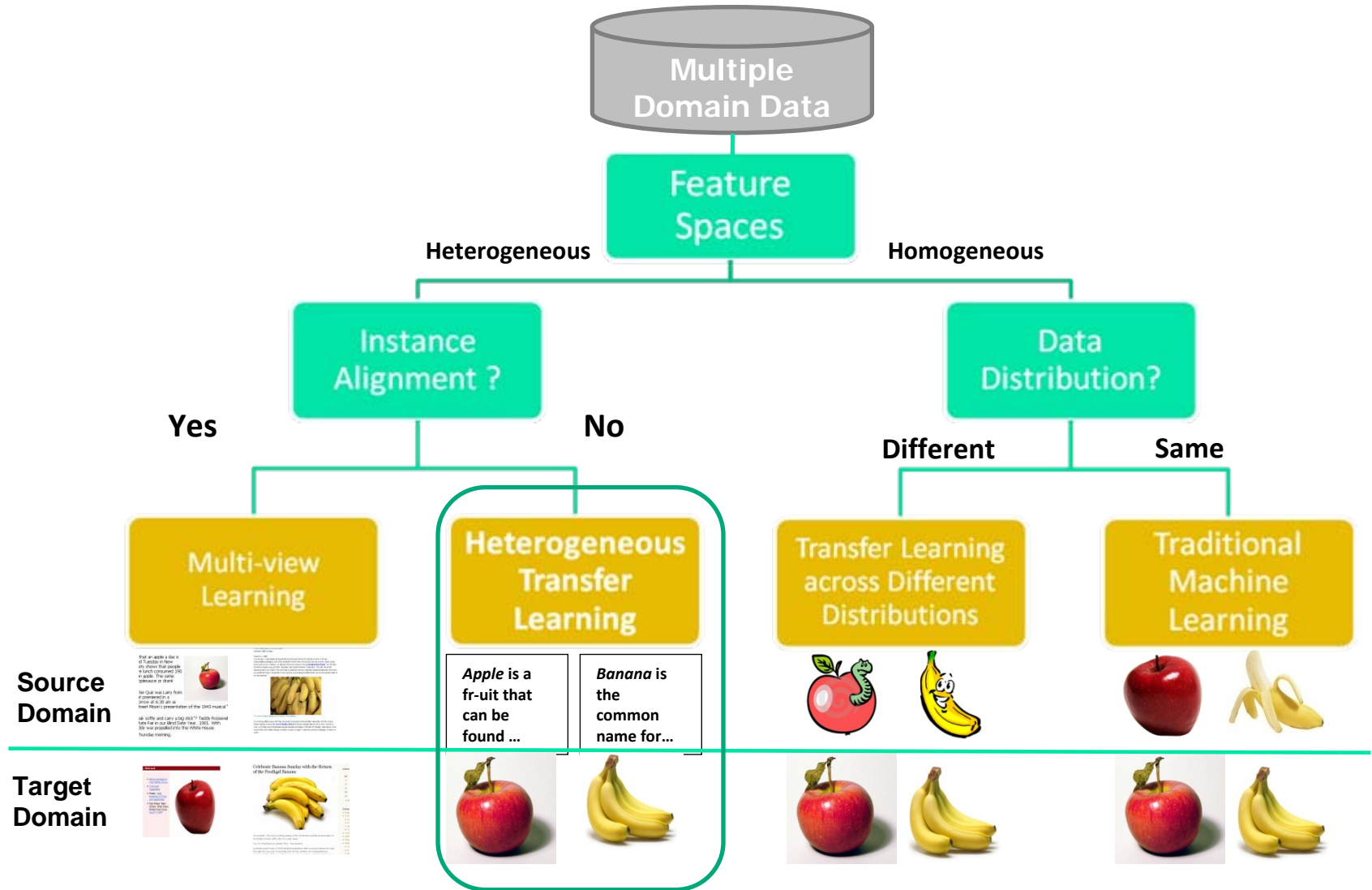(1) **The book is so repetitive that I found myself yelling …. I will definitely not buy another.**

(2) **Do not buy the Shark portable steamer …. Trigger mechanism is defective.**

**Book Domain**          **Kitchen Domain**

# Different Learning Problems
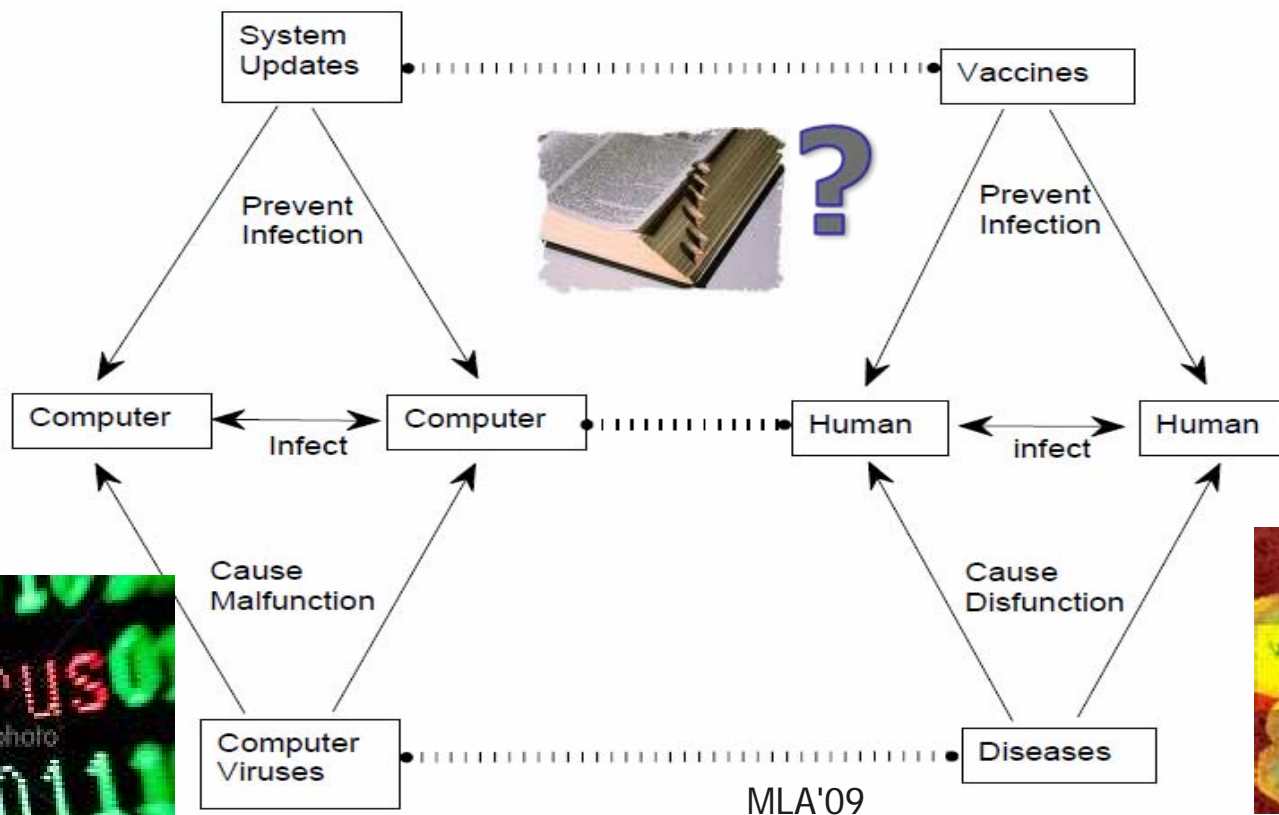
# Outline

- Transfer Learning Basics
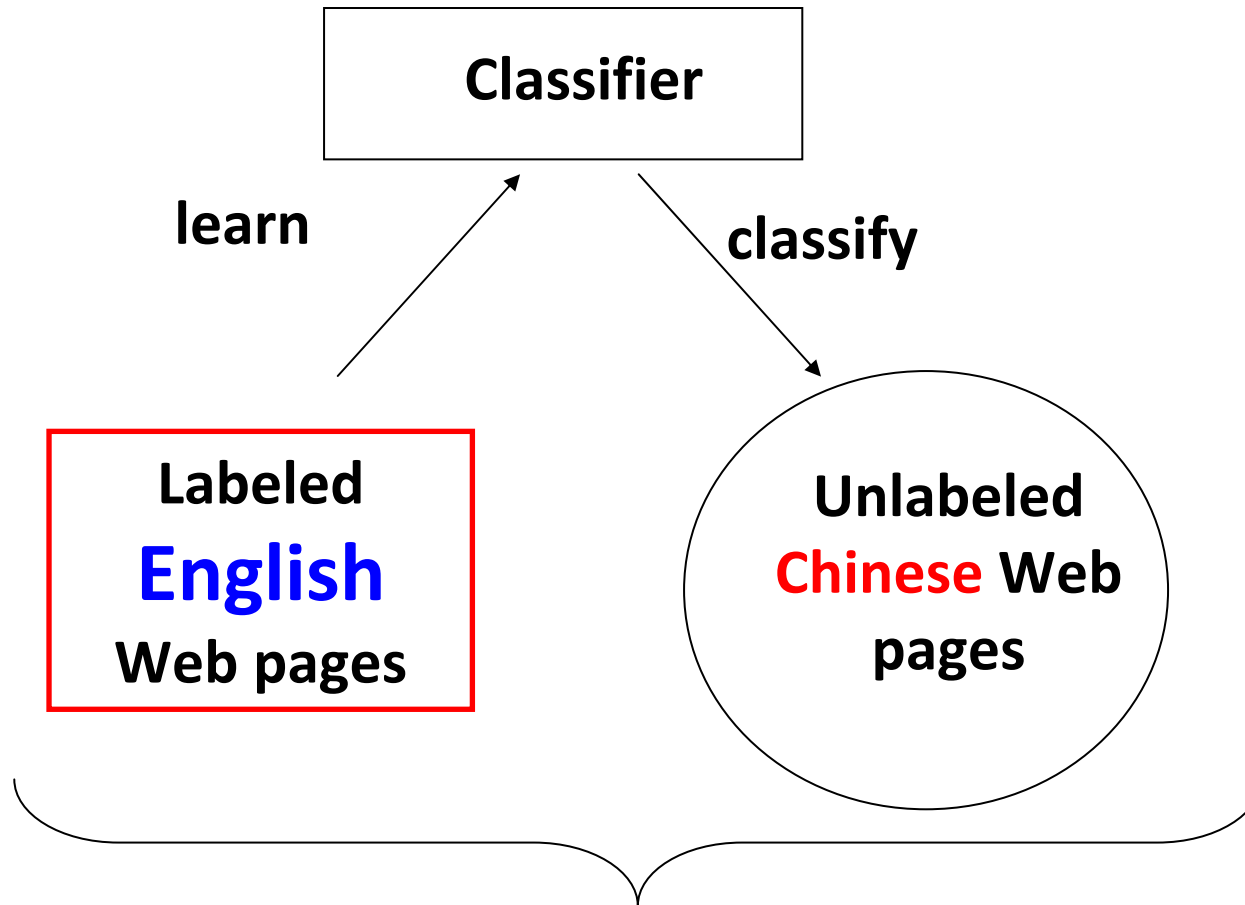- Homogeneous Transfer Learning
- Heterogeneous Transfer Learning
  - With Correspondence
    - Translated Learning (English → Chinese)
    - Text-to-Image Clustering/Classification
  - Without Correspondence
- Future Works

# Correspondence in Transfer Learning

- Mapping between entities or relations
- Probabilistic in nature

# Cross-language Classification



Cross-language Classification

# Heterogeneous Transfer Learning with a Dictionary

[Bel, et al. ECDL 2003]

[Zhu and Wang, ACL 2006]

[Gliozzo and Strapparava ACL 2006]

**DICTIONARY**

**Labeled documents in English (abundant)**

**Translation Error Topic Drift**

**Labeled documents in Chinese (scarce)**

**TASK: Classifying documents in Chinese**
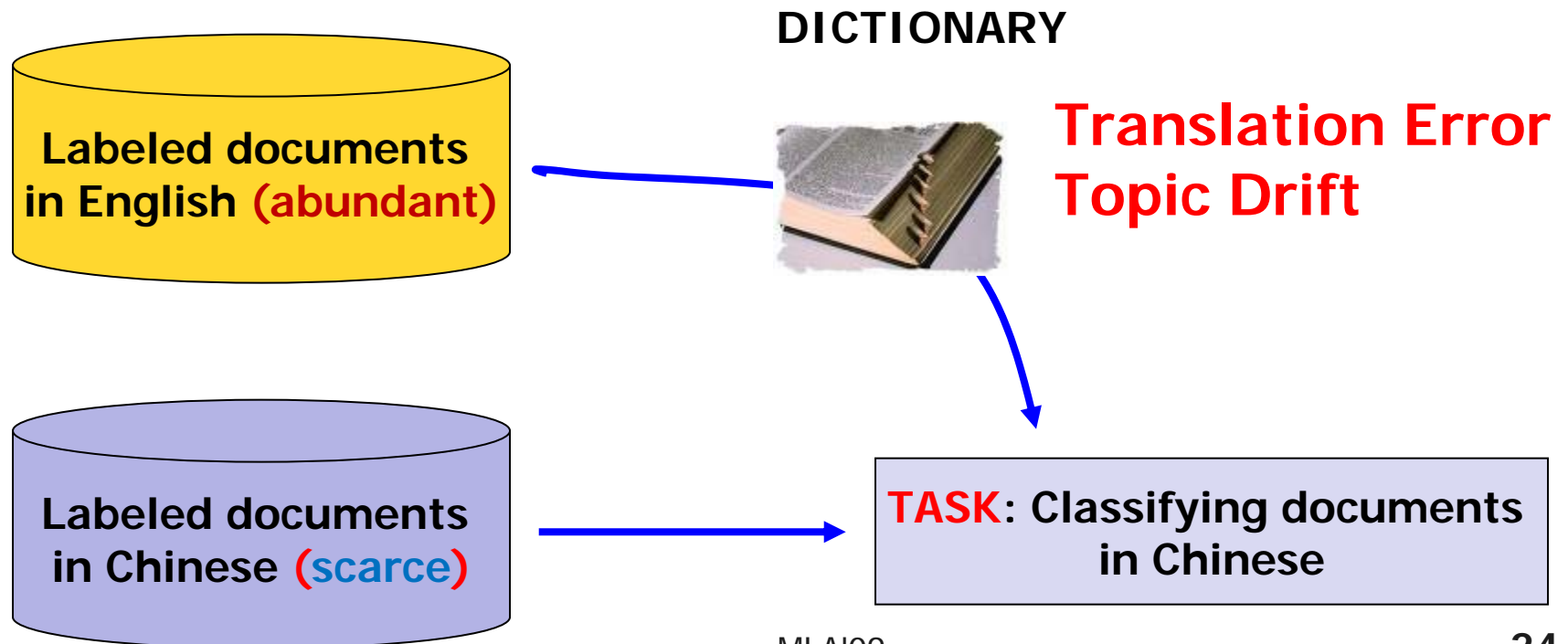
# Topic Drift in Direct Translation

- **Translation Error**
- **Topic Drift**



**Word Features**

in English

in Chinese

News

Games

News

Games
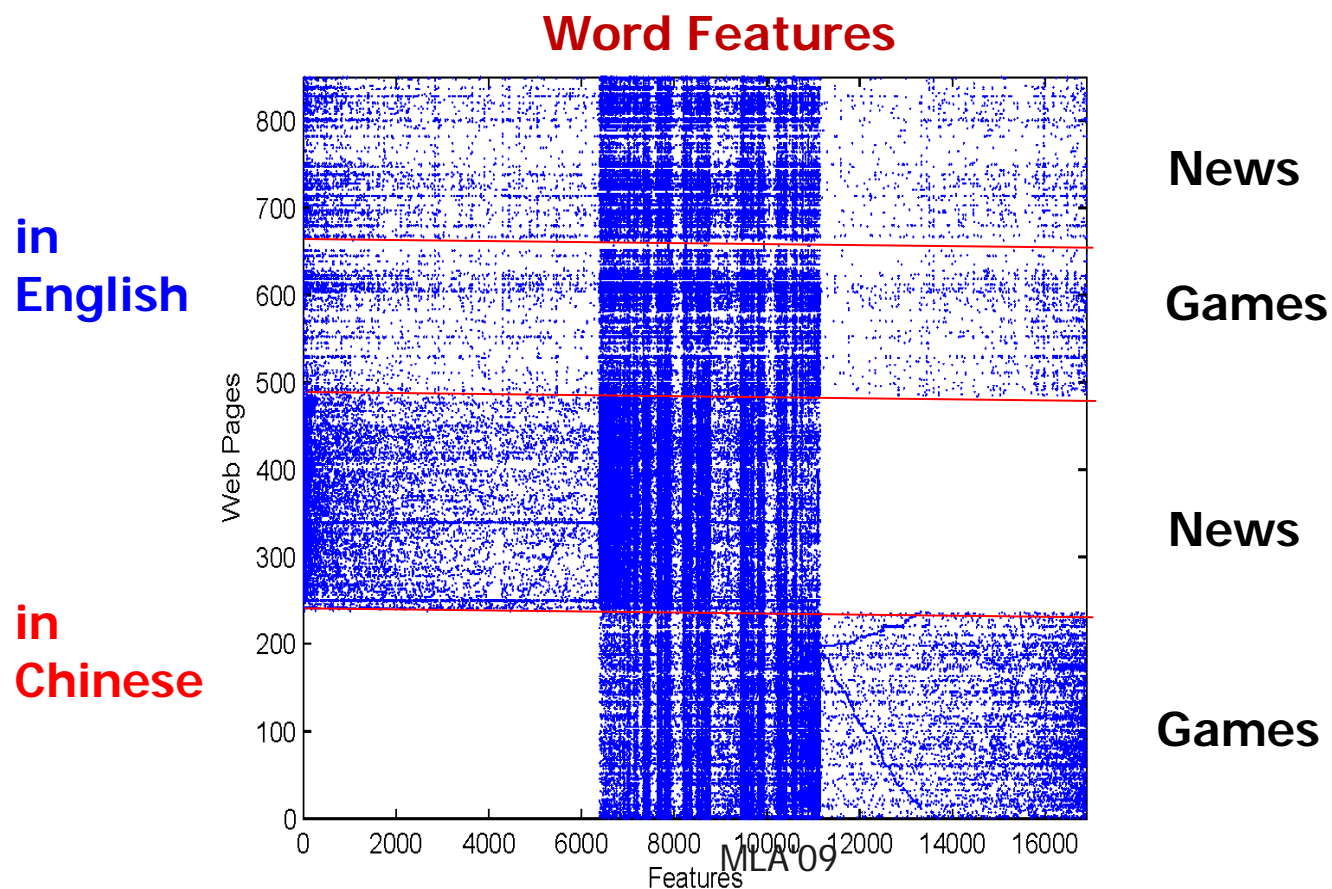
# Improvements: over 15%
[Ling, Xue, Yang et al. WWW2008]

**Domain Adaptation**



Optimize the objective function and give new labels for Chinese Web Pages

Unlabeled Chinese Web Pages

Translator

Unlabeled Chinese Web Pages in English

計算機    互聯網
(计算机)    (互联网)

Labeled English Web Pages

Basic Classifier

New Labels for Chinese Web pages

N times

Output the labels for Chinese Web pages
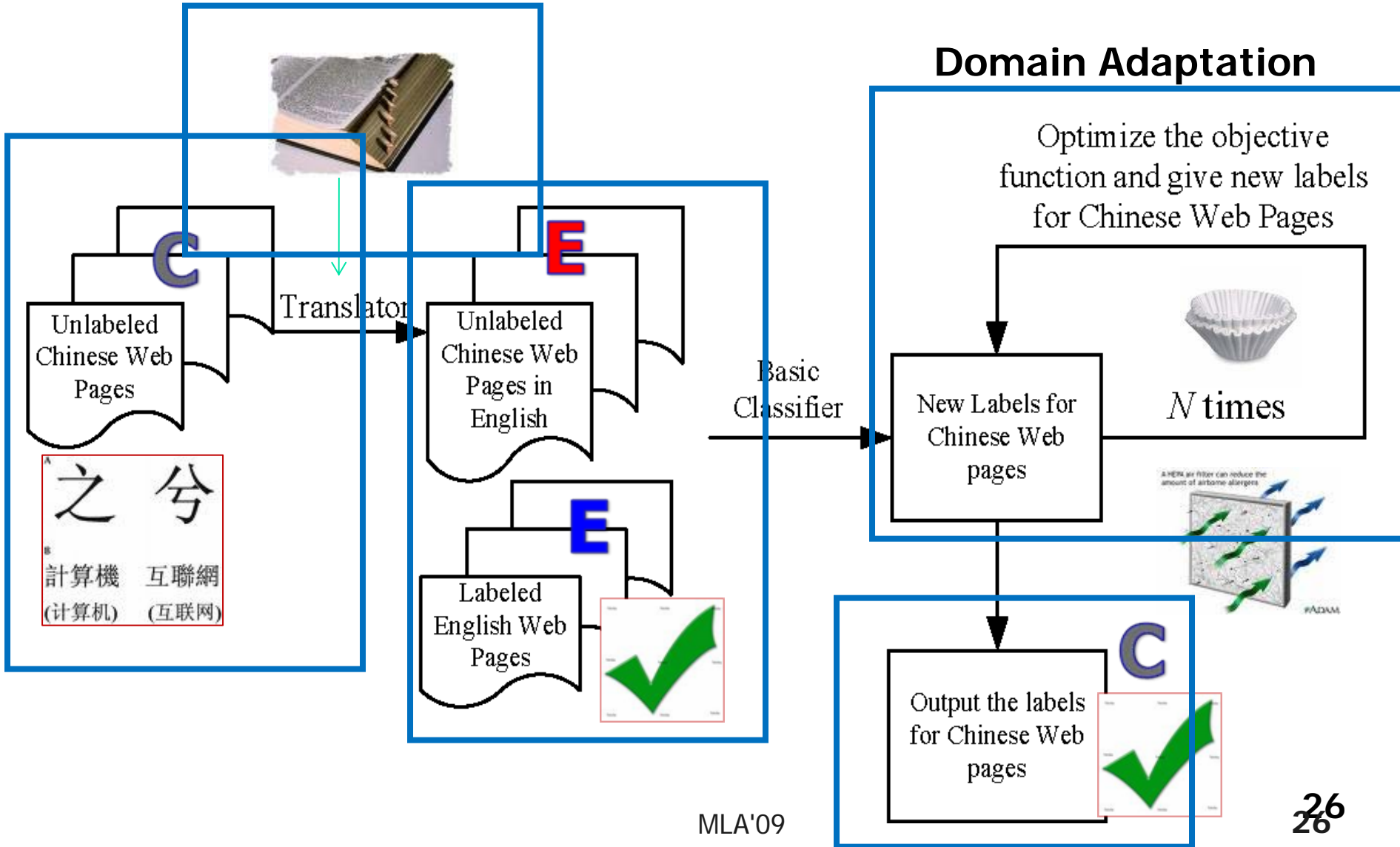
# Text-aided Image Clustering

- Objective: Image clustering

**Apple =**

OR



Apple [ Search images ]

Apple
600 x 637 - 19k - jpg
synergyblog.wordpress.com

Earlier today, **Apple**
2400 x 2400 - 344k - jpeg
operationgadget.com

**Apple's** retail stores
530 x 630 - 36k - jpg
letsgodigital.org

Apple
500 x 555 - 18k - png
cnymultimedia.com

Credit: **Apple Computer**, Inc.
320 x 384 - 66k - jpg
idsa.org

**Apple Computer** has been
402 x 480 - 22k - gif
blorge.com

**Apple computer** made out of
1115 x 871 - 529k - jpg
home.earthlink.net

**Apple** - Fixing-their-shit -
400 x 349 - 14k - jpg
plusvsminus.com

MLA'09

# Adding Auxiliary Text Data



Clustering Model

Search Result Clusters

# Annotated PLSA Model for Clustering $z$

| Caltech 256 Data | Heterogeneous Transfer Learning |
|---|---|
| Average Entropy Improvement | 5.7% |

**SIFT Features**

**From Flickr.com**



**Words from Source Data**

**Topics**

**Image features**

$$P(z|w)$$

$$W$$

$$P(f|z)$$

$$z \rightarrow F$$

$$V$$

$$P(z|v)$$

**Image instances in target data**
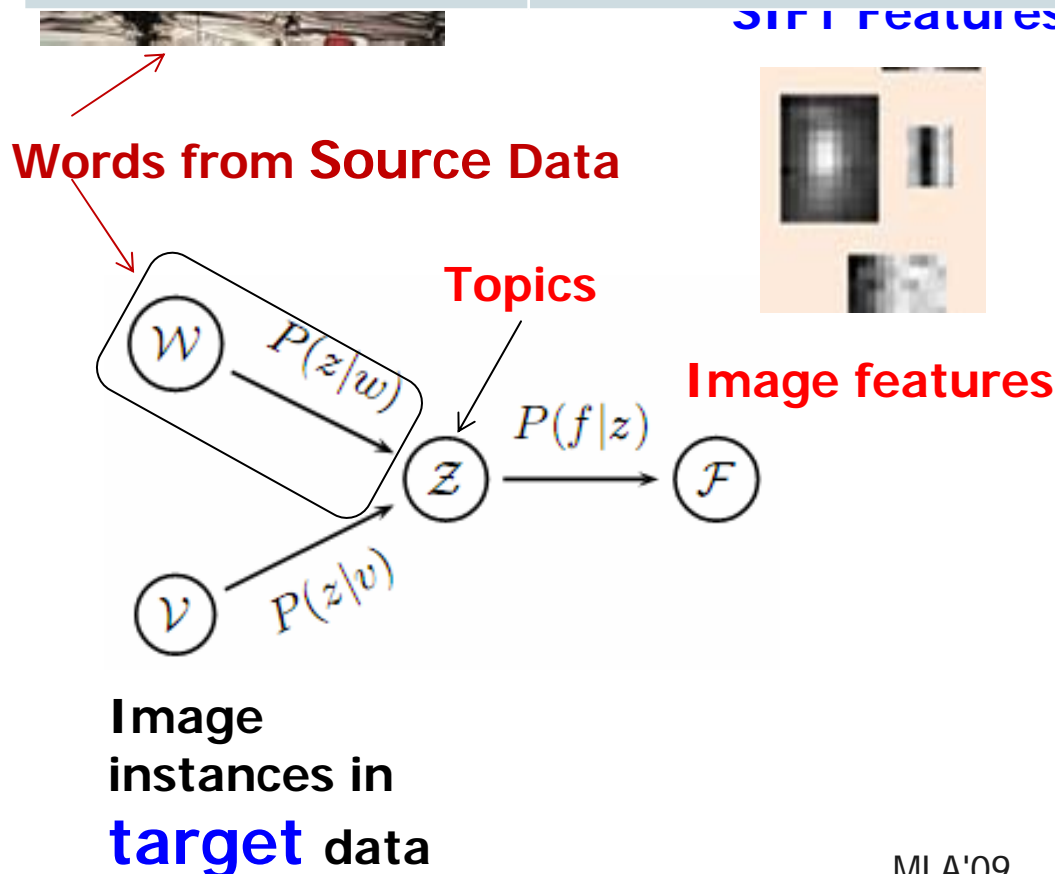
... **Tags**
**Lion**
**Animal**
**Simba**
**Hakuna**
**Matata**
**FlickrBigCats**
...

MLA'09

# Text to Image Classification
## [Dai, Chen, Yang et al. NIPS 2008]



Apple is a fruit. Apple pie is...

Input → **Text Classifier** → Output → **+** / **−**

translating learning models

Input → **Image Classifier** → Output → **+** / **−**
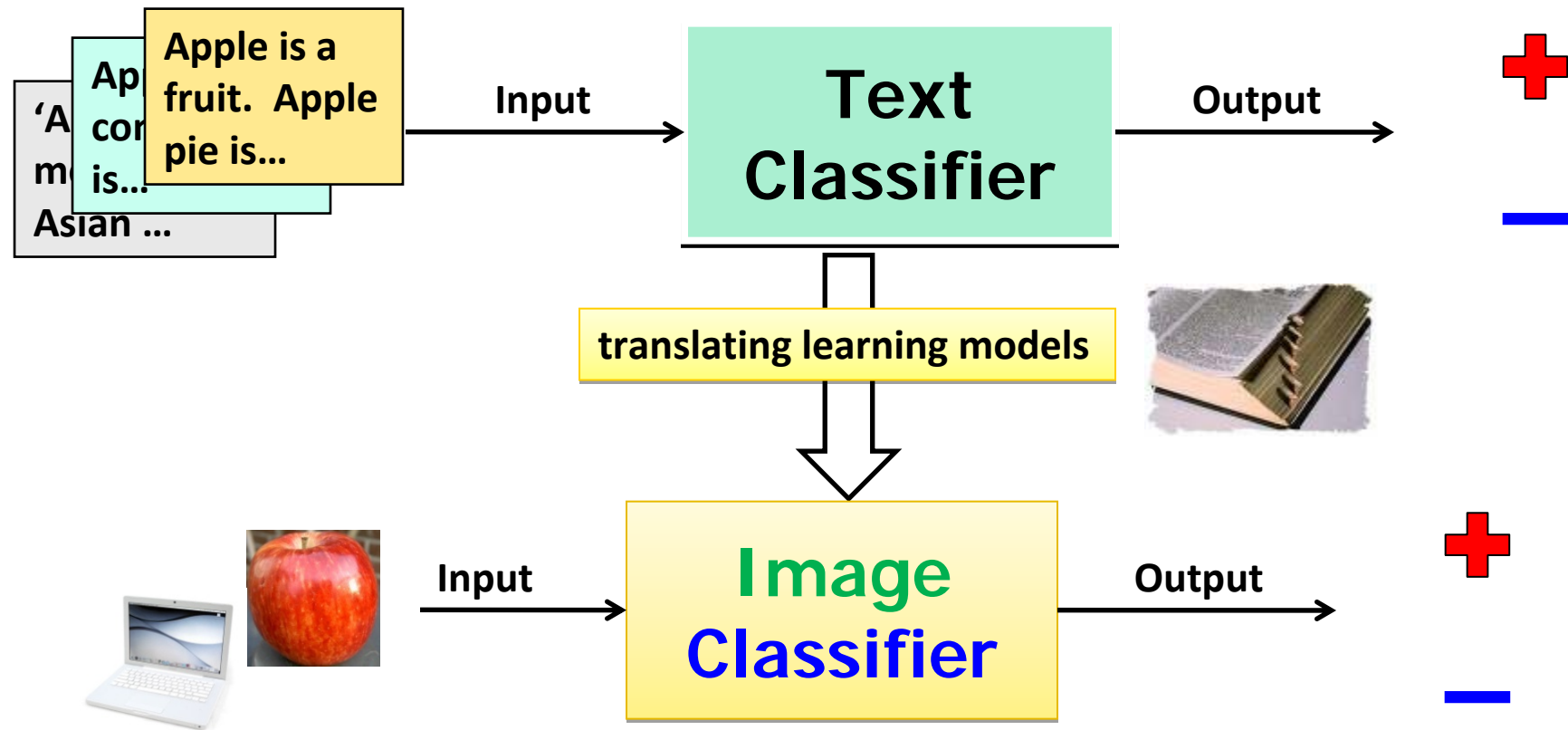
# Heterogeneous Transfer Learning with Correspondence
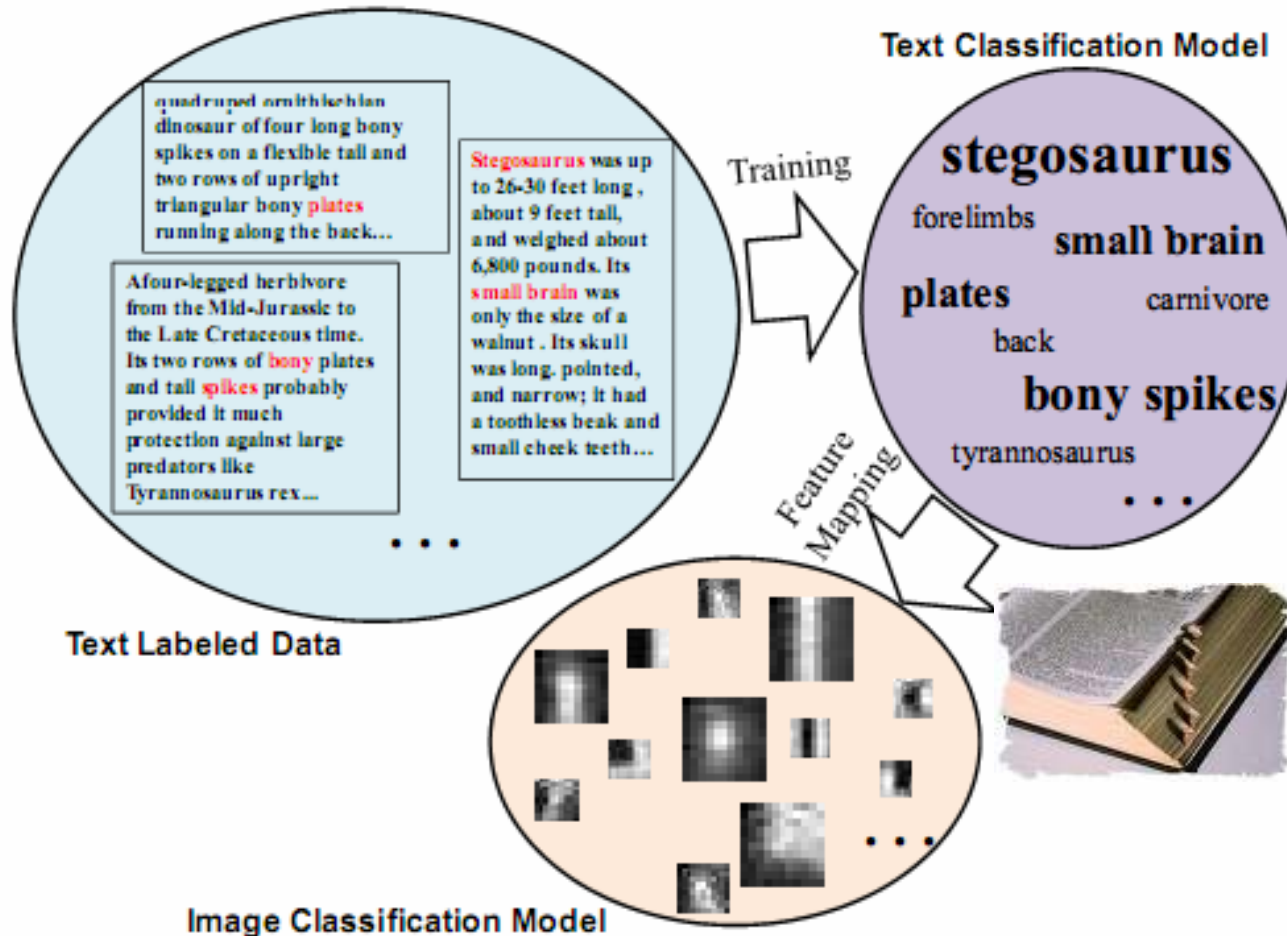
**Log-likelihood:** $h_{\mathrm{TAIC}}(v) = \arg\max_{c \in \mathcal{C}} \log \Pr(c|\mathcal{D}) + \lambda \log \Pr(c|\mathcal{V})$

$$+ \sum_{f \in v} \boxed{\log \Pr(f|c,\mathcal{D})} + \lambda \sum_{f \in v} \log \Pr(f|c,\mathcal{V}).$$

**difficult to estimate!**

$P(f \mid c, D) =$

$\blacksquare$ $\Pr(\;$  $\mid$ ,  $)$

# 14% error reduction
# On Caltech 256 data

$\underbrace{\Pr(f \mid c, w)}_{\text{feature mapping}}$

$\underbrace{\Pr(v \mid w, c)}_{\text{co-occurence data}} = \;\; \bullet \; \Pr(\;$  $\mid$ , APPLE $)$

**Social Web**
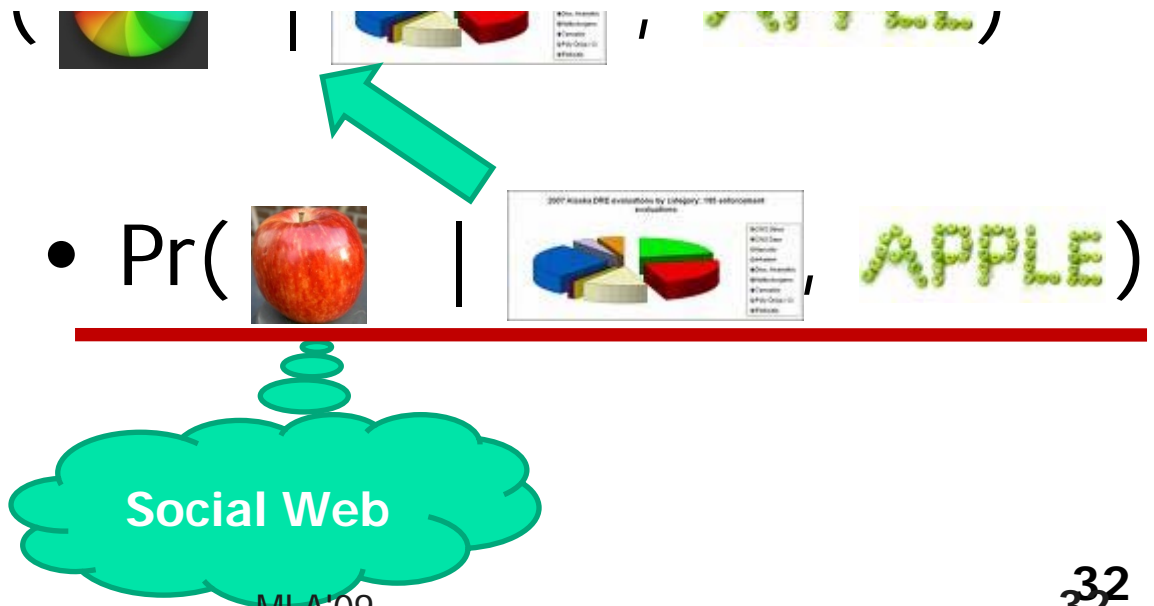
# Outline

- **Homogeneous**
  - Instance Based Transfer
  - Feature Based Transfer
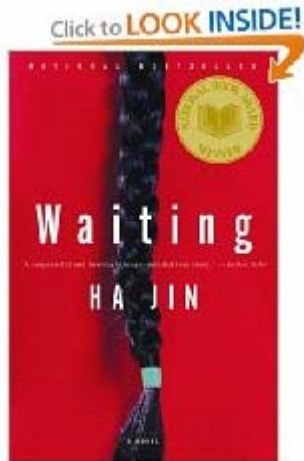- **Heterogeneous (w/ Correspondence)**
- **Heterogeneous Transfer w/out Correspondence**
  - Transfer Learning in Collaborative Filtering
  - Structure-based Transfer
- **Future Works**

# Product Recommendation (Amazon.com)



Click to LOOK INSIDE!
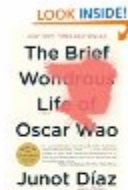
**Customers Who Bought This Item Also Bought**
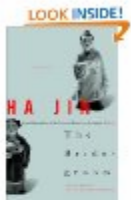
War Trash by Ha Jin
(45) $10.17

The Brief Wondrous Life of Oscar Wao by Junot Díaz
(402) $10.78

The White Tiger: A Novel (Man Booker Prize) by Aravind Adiga
(237) $8.40

The Bridegroom: S by Ha Jin
(27) $1

# Collaborative Filtering: Data Sparseness
(1: don't like; 3: like)

**Products**

**Users**

**Dense (75%)**

| | a | b | c | d | e | f |
|---|---|---|---|---|---|---|
| **1** | ? | 3 | ? | 3 | 2 | 3 |
| **2** | 3 | 1 | 2 | 2 | ? | 1 |
| **3** | 3 | ? | 2 | ? | 3 | 1 |
| **4** | 3 | ? | 1 | 1 | ? | 2 |
| **5** | 2 | 3 | 3 | 3 | 2 | ? |
| **6** | 3 | 2 | ? | 1 | 3 | 2 |

**Nearest neighbor** →

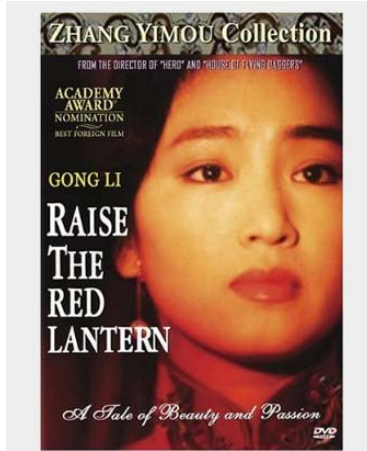| | a | b | c | d | e | f |
|---|---|---|---|---|---|---|
| **2** | 3 | 1 | 2 | 2 | ? | 1 |
| **3** | 3 | ? | 2 | ? | 3 | 1 |
| **1** | ? | 3 | ? | 3 | 2 | 3 |
| **5** | 2 | 3 | 3 | 3 | 2 | ? |
| **4** | 3 | ? | 1 | 1 | ? | 2 |
| **6** | 3 | 2 | ? | 1 | 3 | 2 |

**Overlaps:**
**MORE**

**Similarity:**
**RELIABLE**

# Transfer Learning for Collaborative Filtering?

**IMDB Database**

## Recommendations

If you enjoyed this title, our database also recommends:

**The Good Earth**
IMDb User Rating:
★★★★★★★☆☆☆

**King Lear**
IMDb User Rating:
★★★★★★★☆☆☆

**Big Fish**
IMDb User Rating:
★★★★★★★☆☆☆

**Shi mian mai fu**
IMDb User Rating:
★★★★★★★☆☆☆
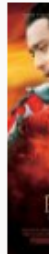
IMD
★★★
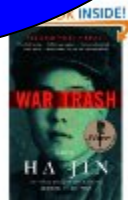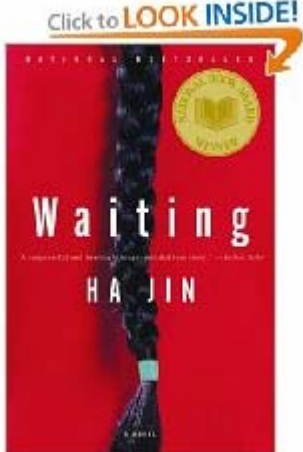
Show more recommendations

**Amazon.com**

Click to LOOK INSIDE!
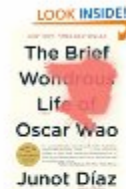
## Customers Who Bought This Item Also Bought

**War Trash** by Ha Jin
★★★★☆ (45) $10.17

**The Brief Wondrous Life of Oscar Wao** by Junot Díaz
★★★★☆ (402) $10.78

**The White Tiger: A Novel (Man Booker Prize)** by Aravind Adiga
★★★☆☆ (237) $8.40

**The Bridegroom: S** by Ha Jin
★★★★☆ (27) $1

MLA'09

# Transfer Learning for Collaborative Filtering [B. Li, Yang, Xue, ICML 2009]

- Users are related in Interests; Items are related in Genre
- How to "Relate" users and items?
    - ALIGN user/item-groups across domains



**MOVIES** (e.g. IMDB)

|     | A       | B       |
|-----|---------|---------|
| I   | 3 3 / 3 3 | 1 1 / 1 1 |
| II  | 2 2 / 2 2 | 3 3 / 3 3 |

**BOOKS** (e.g. Amazon)

|     | A       | B       |
|-----|---------|---------|
| I   | 3 3 / 3 3 | 2 2 / 2 2 |
| II  | 1 1 / 1 1 | 3 3 / 3 3 |

**User/Item Group-Matching**

A ⇔ B    Romance Movies/Books
B ⇔ A    Sci-Fi Movies/Books

I ⇔ II    Girls on IMDB/Amazon
II ⇔ I    Boys on IMDB/Amazon
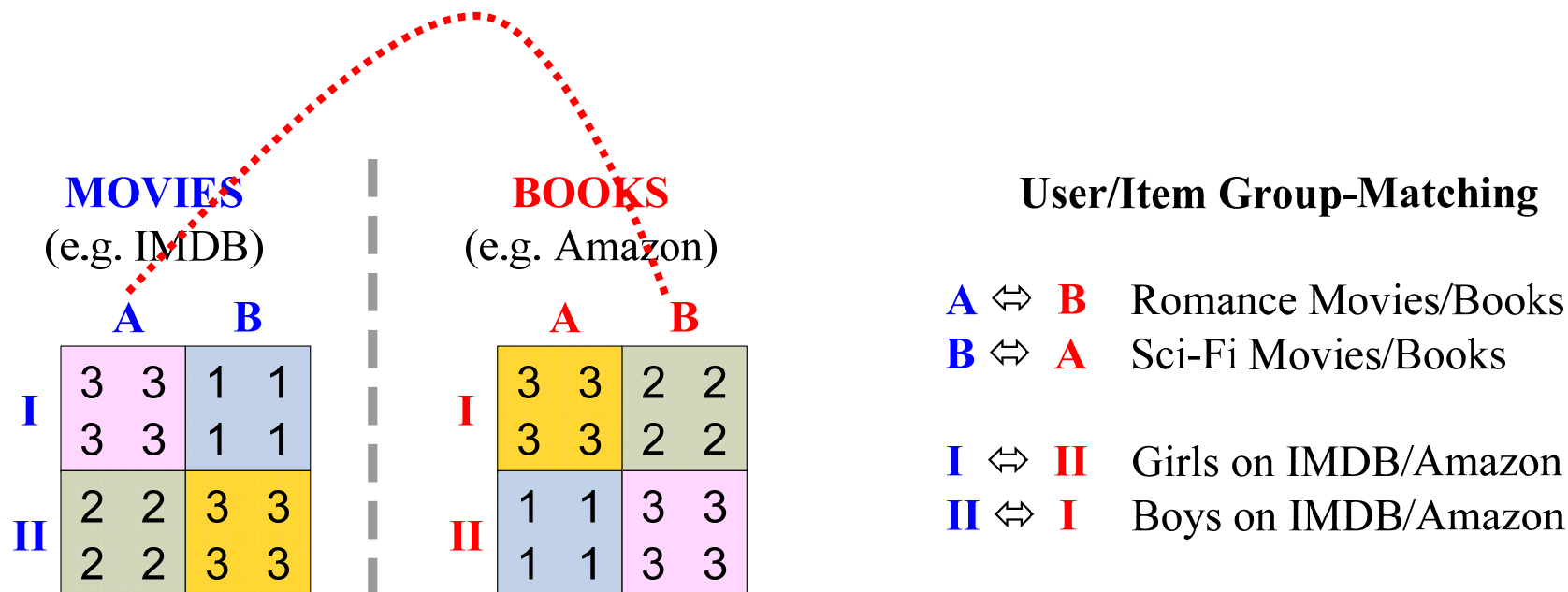
# Codebook based Transfer

|   | a | b | c | d | e |
|---|---|---|---|---|---|
| 1 | **1** | 1 | 3 | 3 | 1 |
| 2 | 3 | 3 | 2 | **2** | 3 |
| 3 | 2 | 2 | **3** | 3 | **2** |
| 4 | 1 | 1 | 3 | **3** | **1** |
| 5 | 1 | **1** | **3** | 3 | 1 |
| 6 | 3 | **3** | 2 | 2 | 3 |
| 7 | **2** | 2 | 3 | 3 | 2 |

**Reconstructed**

|   |   |   |   |
|---|---|---|---|
| 1 | 1 | 0 | 0 |
| 2 | 0 | 1 | 0 |
| 3 | 0 | 0 | 1 |
| 4 | 1 | 0 | 0 |
| 5 | 1 | 0 | 0 |
| 6 | 0 | 1 | 0 |
| 7 | 0 | 0 | 1 |

**User-cluster Indicators**

$\times$

|   | A | B | C |
|---|---|---|---|
| I | 3 | 1 | 2 |
| II | 2 | 3 | 3 |
| III | 3 | 2 | 1 |

**Codebook**

$\times$

| a | b | c | d | e |
|---|---|---|---|---|
| 0 | 0 | 1 | 1 | 0 |
| 1 | 1 | 0 | 0 | 1 |
| 0 | 0 | 0 | 0 | 0 |

**Item-cluster Indicators**

Dat                                                                              ethods
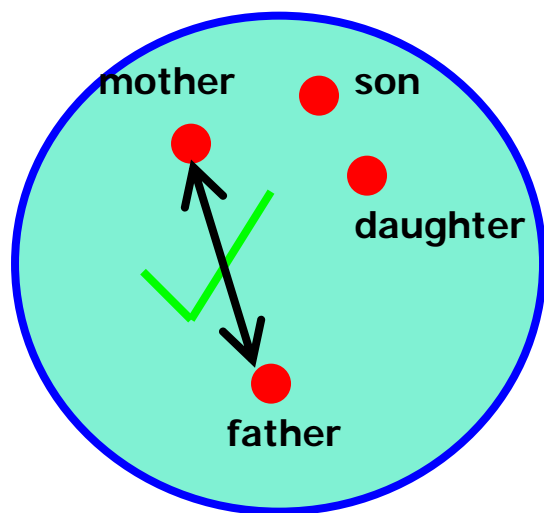- Ea                                                                    ation Coefficients
- Bo                                                                    er-based Smoothing
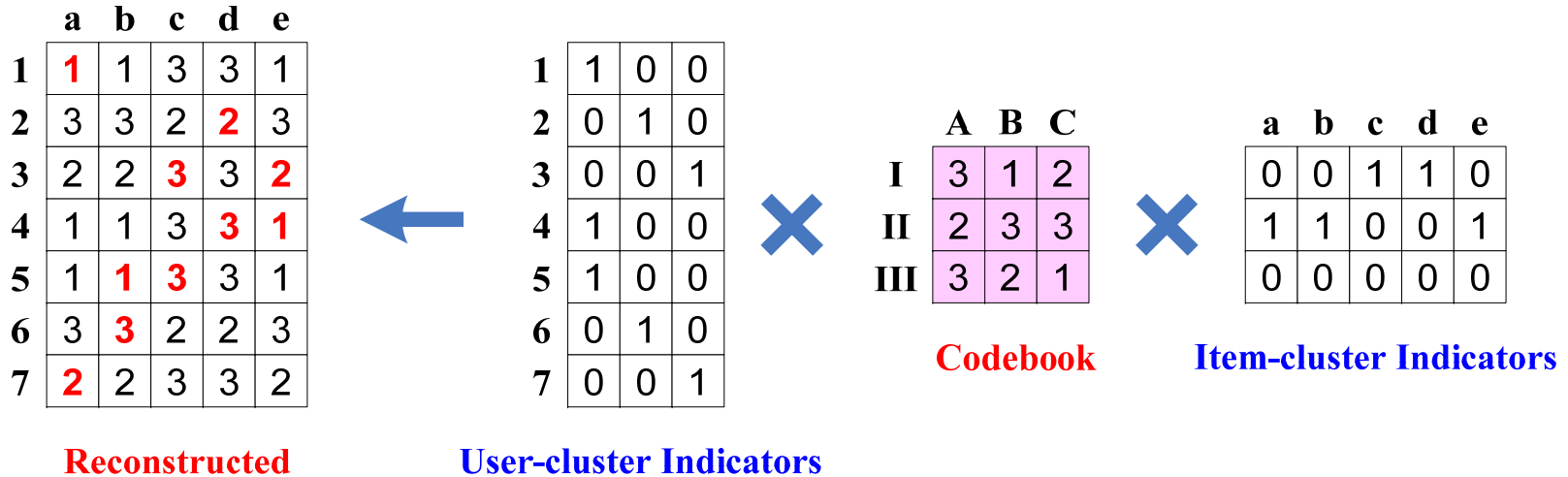
$$\min_{\substack{U_{tgt}\in\{0,1\}^{p\times k} \\ V_{tgt}\in\{0,1\}^{q\times l}}} \left\| [X_{tgt} - U_{tgt}\mathbf{B}V_{tgt}^{\top}] \circ W \right\|_F^2$$

$$\text{s.t.} \quad U_{tgt}\mathbf{1} = \mathbf{1}, V_{tgt}\mathbf{1} = \mathbf{1}$$

- Codebook Transfer (CBT)
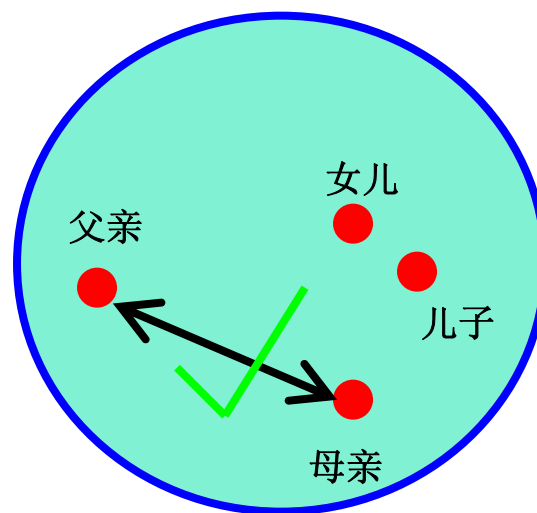
Result: 5-10% improvement

# Heterogeneous Transfer Learning **without** Correspondence [H. Wang and Yang 2009]

*Goal:*

- *Learn a correspondence structure between domains*
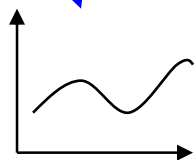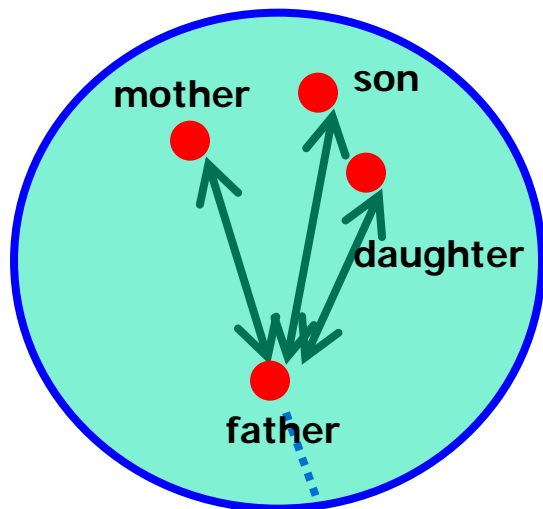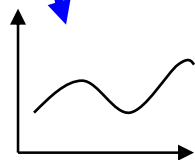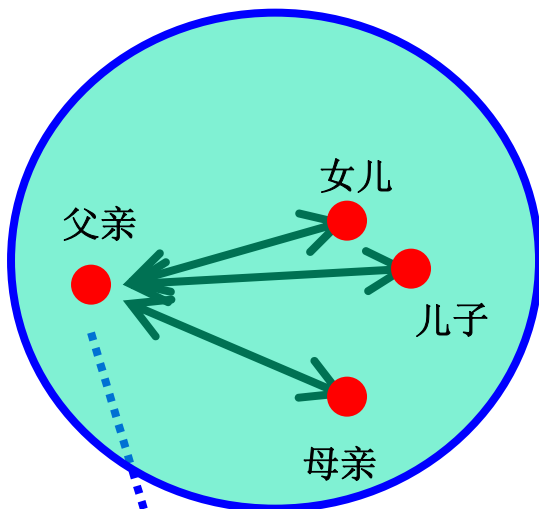- *Use the correspondence to transfer knowledge*



**English**

**Chinese (**汉语）

# Heterogeneous Transfer Learning without correspondence

[Dekang Lin, 'An Information-theoretic Dfn of Similarity', ICML 1998]

**English**

mother
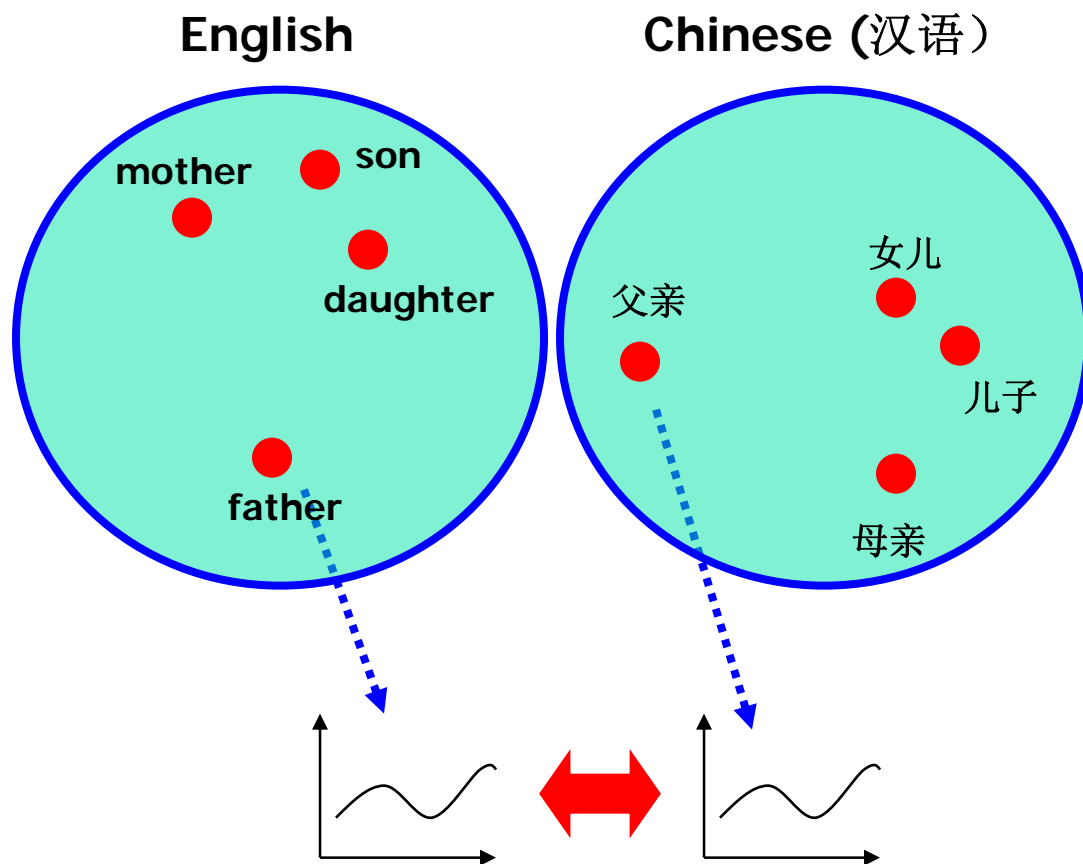son
daughter
father

**Chinese (汉语）**

父亲
女儿
儿子
母亲

**STEP 1:**

• Compare each entity with all others in the same domain.

• Encode each entity by distribution

# Heterogeneous Transfer Learning without correspondence

**English**

**Chinese (汉语）**

mother

son

daughter

father

父亲

女儿

儿子

母亲

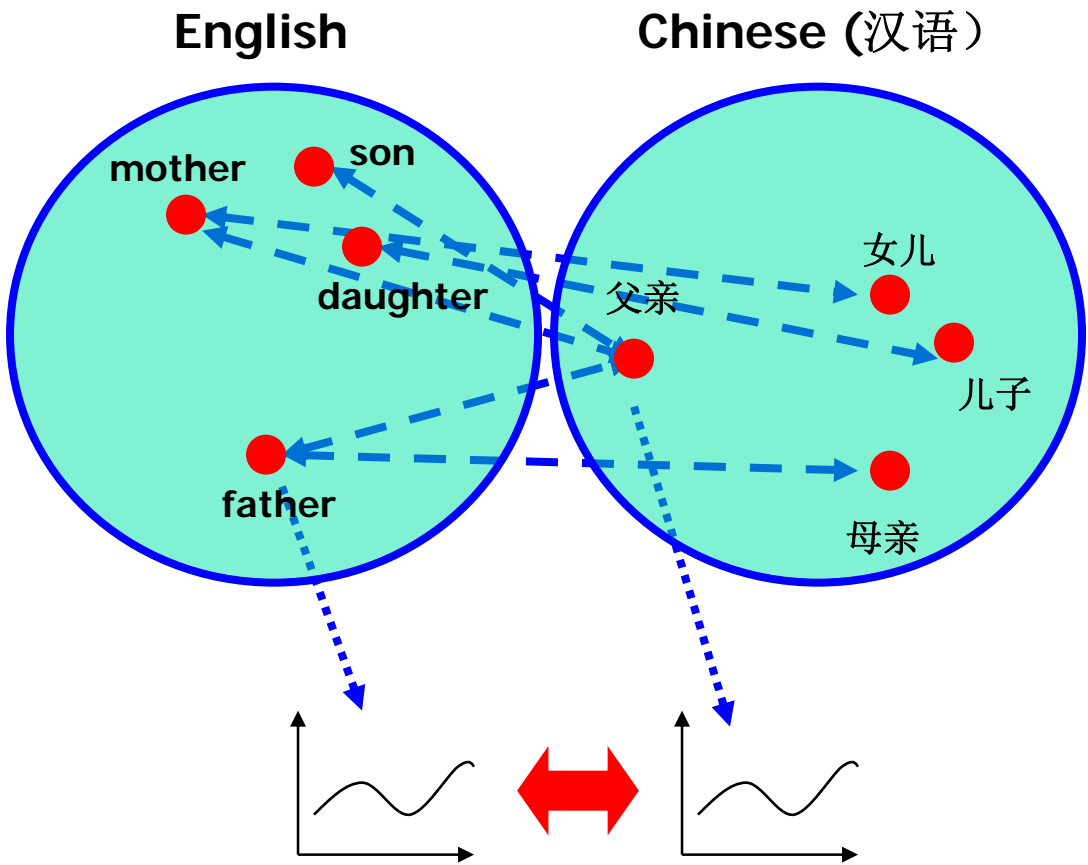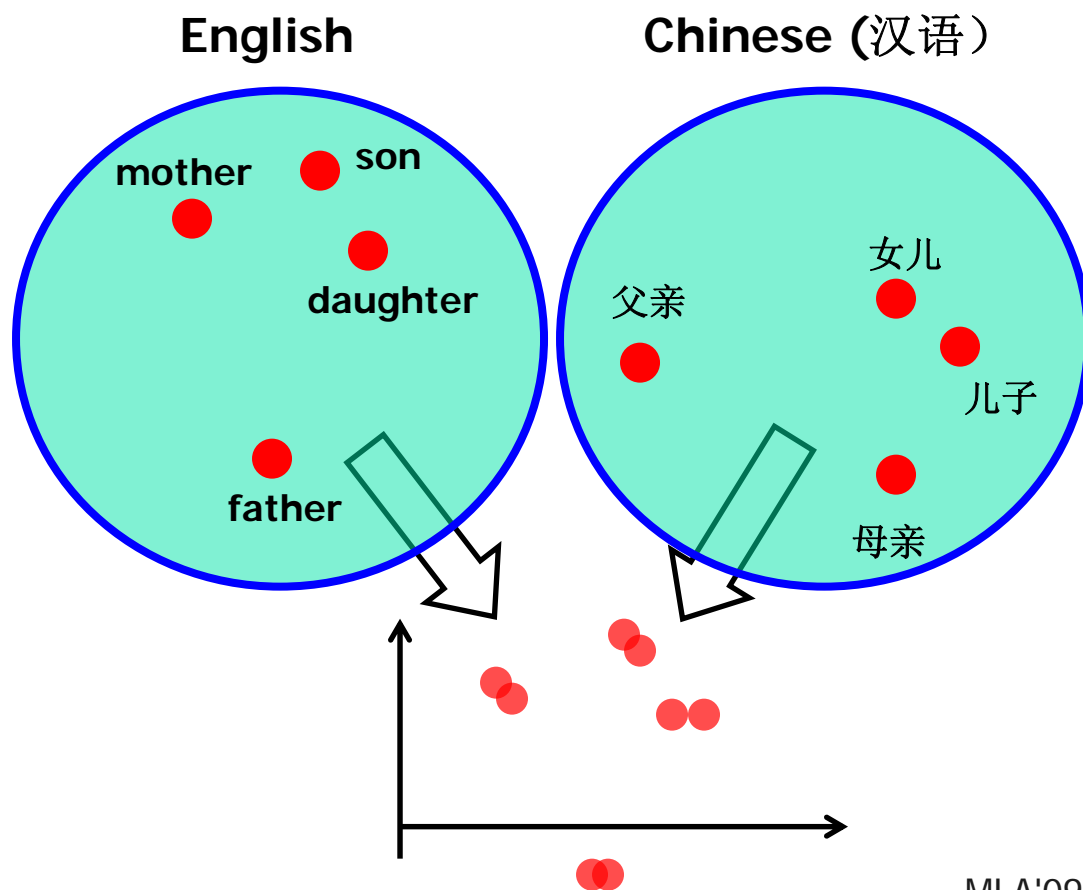**STEP 2:**

**• Compare two distributions in order to measure their *relatedness***

# Heterogeneous Transfer Learning without Correspondence



**English**

**Chinese (**汉语）

mother
son
daughter
father

女儿
父亲
儿子
母亲

STEP 3:

**Build a bipartite graph across the domains,**

# Heterogeneous Transfer Learning without correspondence

**English**

**Chinese (汉语）**

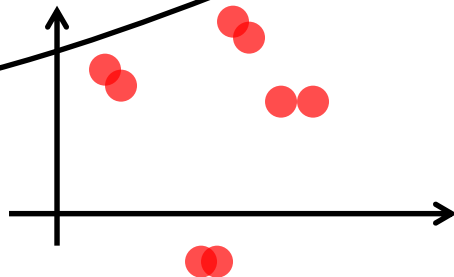mother  son

daughter

father

父亲

女儿

儿子

母亲

**STEP 4:**

**Align the two domains in a common latent space by spectral analysis methods.**

# Heterogeneous Transfer Learning without correspondence



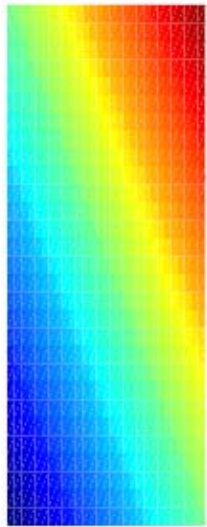**English**   **Chinese (**汉语）

mother
son
daughter
father

女儿
父亲
儿子
母亲
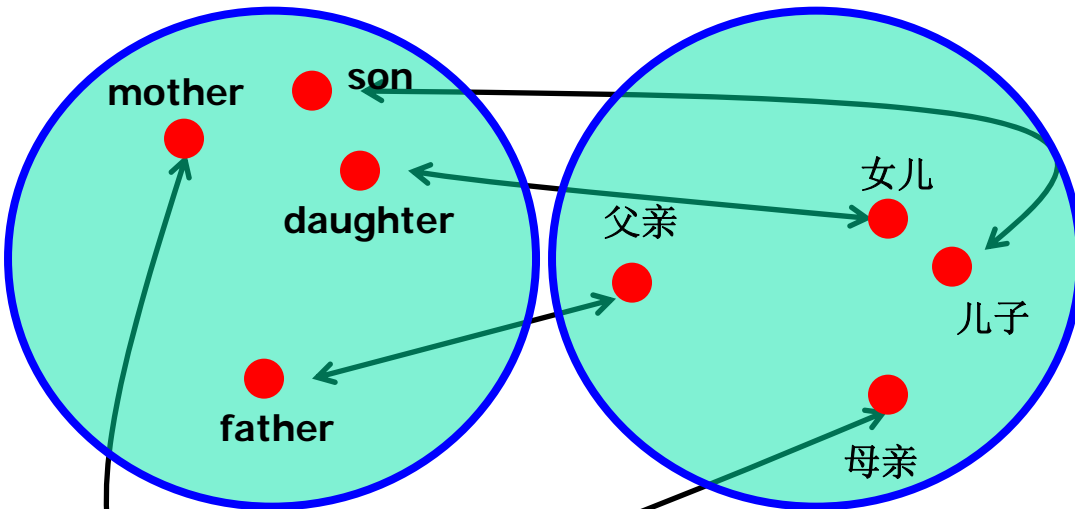
**STEP 5**

**Finally the "common parts" is used for knowledge transfer…**

# Heterogeneous Transfer Learning without correspondence



(a) Ground truth geodesic structure and color coding
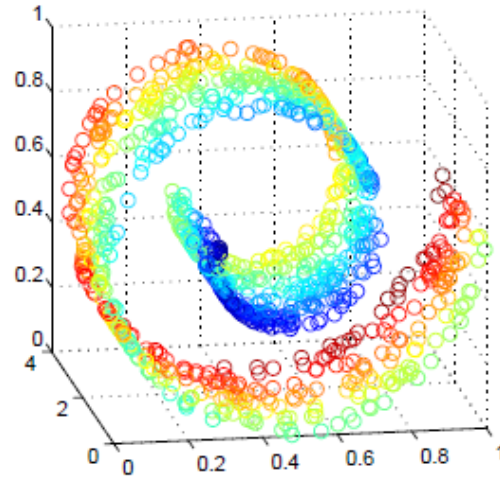
(b) The Swiss roll
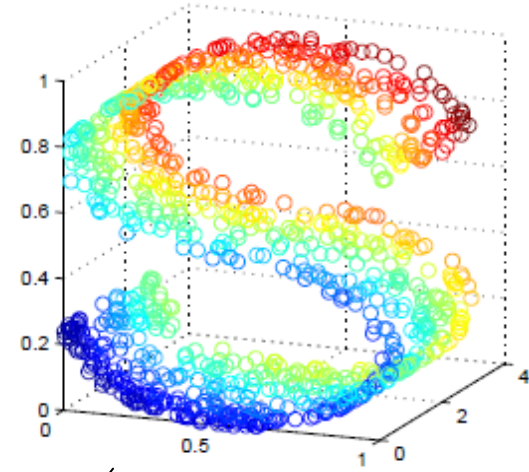
(d) The S curve
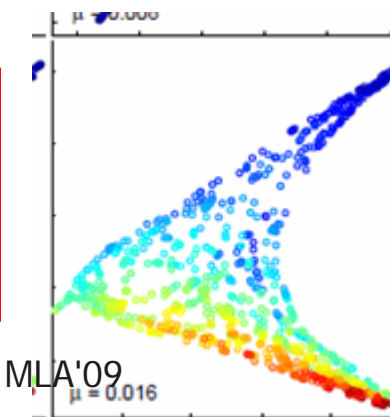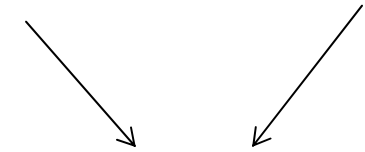
μ = 0.016

# Conclusions and Future Work

- # Homogeneous Transfer Learning

- # Heterogeneous Transfer Learning

  - Feature spaces and distributions are different
  - Methods
    - Known correspondence: Text-based Image Classification/Clustering,
    - Unknown correspondence: Alignment, global structural correspondence

- # Future

  - Negative Transfer
  - Multiple source domains [Gao, Fan, Jiang, Han KDD08] [Luo et al. CIKM 08]
  - Scaling up

# Future: Negative Transfer
## Credit: Dai, Wenyuan
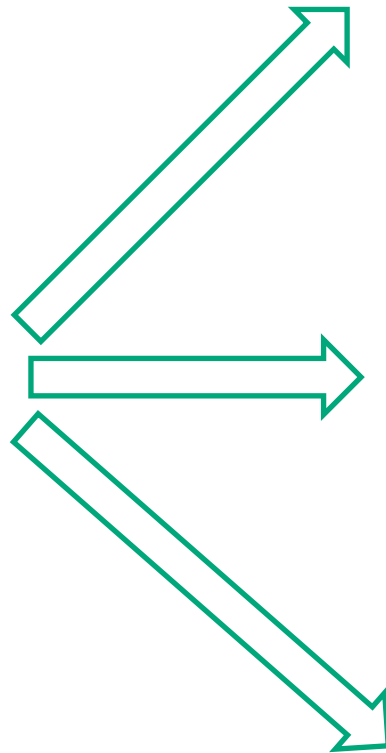


**Helpful**: positive transfer

**Harmful:** negative transfer

**Neutral**: zero transfer

# Future: Negative Transfer

- "To Transfer or Not to Transfer"
    - Rosenstein, Marx, Kaelbling and Dietterich
    - Inductive Transfer Workshop, NIPS 2005.  (Task: meeting invitation and acceptance)
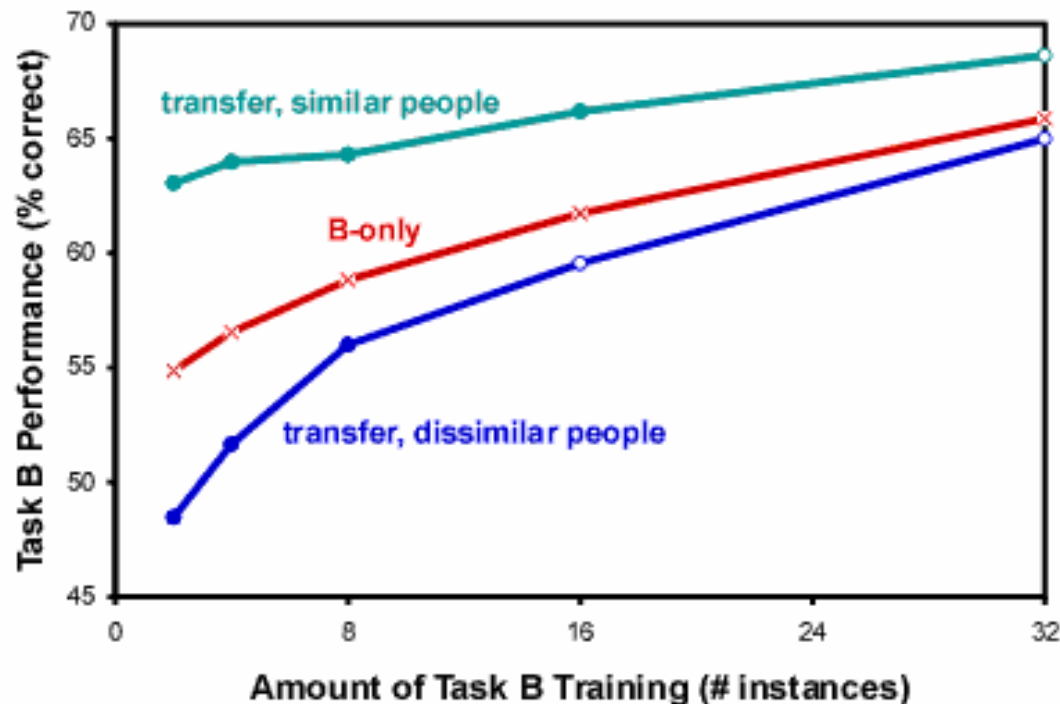
# Acknowledgement

- ## HKUST:
  - Sinno J. Pan, Huayan Wang, Bin Cao, Evan Wei Xiang, Derek Hao Hu, Nathan Nan Liu, Vincent Wenchen Zheng

- ## Shanghai Jiaotong University:
  - Wenyuan Dai, Guirong Xue, Yuqiang Chen, Prof. Yong Yu, Xiao Ling, Ou Jin.

- ## Visiting Students
  - Bin Li (Fudan U.), Xiaoxiao Shi (Zhong Shan U.),

# References

- ## Invited Article at ACL/IJCNLP 2009
  - Qiang Yang,Yuqiang Chen Gui-Rong Xue Wenyuan Dai Yong Yu, <u>Heterogeneous Transfer Learning for Image Clustering via the Social Web</u>. In Proceedings of the ACL/IJCNLP 2009, Singapore, Aug 2-7, 2009. Pages 1-9.

- ## Survey Article
  - Sinno J. Pan and Qiang Yang. *A Survey on Transfer Learning*. IEEE TKDE 2009 (to appear). http://www.cse.ust.hk/~sinnopan/SurveyTL.htm

# References: Surveys

- J. Jiang *A Literature Survey on Domain Adaptation of Statistical Classifiers.* *http://www.mysmu.edu/faculty/jingjiang/*

- L. Torrey and J. Shavlik. *Transfer Learning* http://pages.cs.wisc.edu/~ltorrey/papers/torrey_chapter09.pdf

# References: Domain Adaptation

- [Daume III & Marcu, JAIR 2006] Hal Daumé III, Daniel Marcu: *Domain Adaptation for Statistical Classifiers*. J. Artif. Intell. Res. (JAIR) 26: 101-126 (2006)

- [Blitzer et al. ACL 2007] John Blitzer, Mark Dredze, and Fernando Pereira. Biographies, Bollywood, Boom-boxes, and Blenders: *Domain Adaptation for Sentiment Classification*.

- [Ando and Zhang, JMLR 2005] Rie K. Ando and Tong Zhang. *A Framework for Learning Predictive Structures from Multiple Tasks and Unlabeled Data*. JMLR, 6:1817-1853, 2005.

# References: Domain Adaptation

- [Pan et al. AAAI 08] Sinno Jialin Pan, James T. Kwok and Qiang Yang. *Transfer Learning via Dimensionality Reduction.* In Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI-08). Chicago, Illinois, USA. July 13-17, 2008. Pages 677-682.

- [Pan et al. IJCAI 09] Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok and Qiang Yang. Domain Adaptation via Transfer Component Analysis. In (IJCAI-09). Pasadena, California, USA. July 11-17, 2009.

- [Jing Jiang & Chengxiang Zhai. ACL 07] Jing Jiang, Chengxiang Zhai: *Instance Weighting for Domain Adaptation in NLP*. ACL 2007

- [Mansour et al. NIPS 08] Yishay Mansour, Mehryar Mohri, Afshin Rostamizadeh: *Domain Adaptation with Multiple Sources*. NIPS 2008: 1041-1048

# References: Transfer Learning Applications

- [Davis and Domingos 2009] Jesse Davis and Pedro Domingos. *Deep transfer via second-order Markov logic.* ICML 2009. 217-224. 2009.

- L. Torrey, J. Shavlik, S. Natarajan, P. Kuppili & T. Walker (2008). Transfer in Reinforcement Learning via Markov Logic Networks. *AAAI'08 Workshop on Transfer Learning for Complex Tasks*, Chicago, IL. 2008

- Lilyana Mihalkova, Tuyen Huynh, Raymond J. Mooney Mapping and Revising Markov Logic Networks for Transfer Learning. In *Proceedings of the 22nd AAAI Conference on Artificial Intelligence (AAAI-2007)*, Vancouver, BC, pp. 608-614, July 2007.

- Lilyana Mihalkova and Raymond Mooney . Transfer Learning from Minimal Target Data by Mapping across Relational Domains. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI-09)* , Pasadena, CA, July 2009.

# References: Transfer Learning

- A List of Transfer Learning Papers:

  - http://www.cs.berkeley.edu/~russell/classes/cs294/f05/all-papers.html

  - http://www.cse.ust.hk/~sinnopan/conferenceTL.htm

- [Wu and Dietterich ICML-04] Wu, P. and Dietterich, T. G. (2004). Improving svm accuracy by training on auxiliary data sources. In Proc. ICML, pages 871-878.

- [Dai, Xue, Yang et al. ICML 2007] Wenyuan Dai, Qiang Yang, Gui-Rong Xue and Yong Yu. Boosting for Transfer Learning. In ICML'07 Oregon, USA, June 20-24, 2007. 193 - 200

- [Dai, Xue, Yang et al. KDD  2007] Wenyuan Dai, Gui-Rong Xue, Qiang Yang and Yong Yu. Co-clustering based Classification for Out-of-domain Documents. In ACM SIGKDD '07, Aug 12-15, 2007. Pages 210-219

# References: Heterogeneous Transfer Learning

- [Dai, Chen, Yang et al. NIPS 2008] Wenyuan Dai, Yuqiang Chen, Gui-Rong Xue, Qiang Yang, and Yong Yu. <u>Translated Learning</u>. NIPS **2008**, 2008.

- [Ling et al. WWW 2008] Xiao Ling, Gui-Rong Xue, Wenyuan Dai, Yun Jiang, Qiang Yang, Yong Yu: *<u>Can Chinese web pages be classified with English data source?</u>* WWW 2008: 969-978

- [Li, Yang et al. ICML 2009] Bin Li, Qiang Yang and Xiangyang Xue. <u>Transfer Learning for Collaborative Filtering via a Rating-Matrix Generative Model.</u> In Proceedings of the 26th International Conference on Machine Learning (**ICML '09**), Montreal, Canada, 2009.

# References: Heterogeneous

- [Gao et al. KDD 09] Jing Gao, Wei Fan, Yizhou Sun, Jiawei Han: *Heterogeneous source consensus learning via decision propagation and negotiation*. KDD 2009: 339-348

- [Xing et al. UAI 2005] E.P. Xing, R. Yan and A. G. Hauptmann, Mining Associated Text and Images with Dual-Wing Harmoniums. Uncertainty in Artificial Intelligence, UAI 2005.

- [Luo et al. CIKM 08] Ping Luo, Fuzhen Zhuang, Hui Xiong, Yuhong Xiong, Qing He: *Transfer learning from multiple source domains via consensus regularization*. CIKM 2008: 103-112

# References: Different Classes Between Domains

- *[Shi et al. ECML 2009] Xiaoxiao Shi,Wei Fan, Qiang Yang, Jiangtao Ren.* <u>Relaxed Transfer of Different Classes via Spectral Partition</u> *European Conference on Machine Learning (ECML) 2009.*

- Dou Shen, Jian-Tao Sun, Qiang Yang, Zheng Chen: <u>Building bridges for web query classification</u>. SIGIR 2006: 131-138

# References: Negative Transfer

- [Rosenstein et al. NIPS 2005] Michael T. Rosenstein, Zvika Marx, Leslie Pack Kaelbling & Thomas G. Dietterich. <u>To Transfer or Not To Transfer</u>. NIPS 2005 Workshop on Inductive Transfer.  2005.

# References: Cross-language Applications

- [Zhu and Wang, ACL 2006] Jiang Zhu, Haifeng Wang: The Effect of Translation Quality in MT-Based Cross-Language Information Retrieval. ACL 2006

- [Gliozzo and Strapparava ACL 2006] Alfio Massimiliano Gliozzo, Carlo Strapparava: Exploiting Comparable Corpora and Bilingual Dictionaries for Cross-Language Text Categorization. ACL 2006

- [Olsson et al. SIGIR 05] J. Scott Olsson, Douglas W. Oard, Jan Hajic: Cross-language text classification. SIGIR 2005: 645-646.

# References: Cross-language Applications

- [Gliozzo et al. 2005] A. Gliozzo and C. Strapparava. Cross language Text Categorization by acquiring Multilingual Domain Models from Comparable Corpora. Proc. of the ACL Workshop on Building and Using Parallel Texts.  2005

- [Bel et al. ECDL 2003] N. Bel, C. Koster, and M. Villegas. Cross-lingual text categorization. In *ECDL, 2003.*

- [Brown et al. 1993] Peter F. Brown, J. Cocke, Stephen A. Della Pietra, Vincent J. Della Pietra, F. Jelinek, Robert L. Mercer, and P.S. Roossin. 1990. *A statistical approach to machine translation*. Computational Linguistics, 16:79-85.