# Ranking with Multiple Hyperplanes

Tao Qin et al.
Microsoft Research Asia
SIGIR 2007

# Problem Definition

- Set of objects
- Each object has a rank
- Order objects according to rank

# Problem Definition

- Set of objects
- Each object has a rank
- Order objects according to rank

- Don’t know ranks in advance

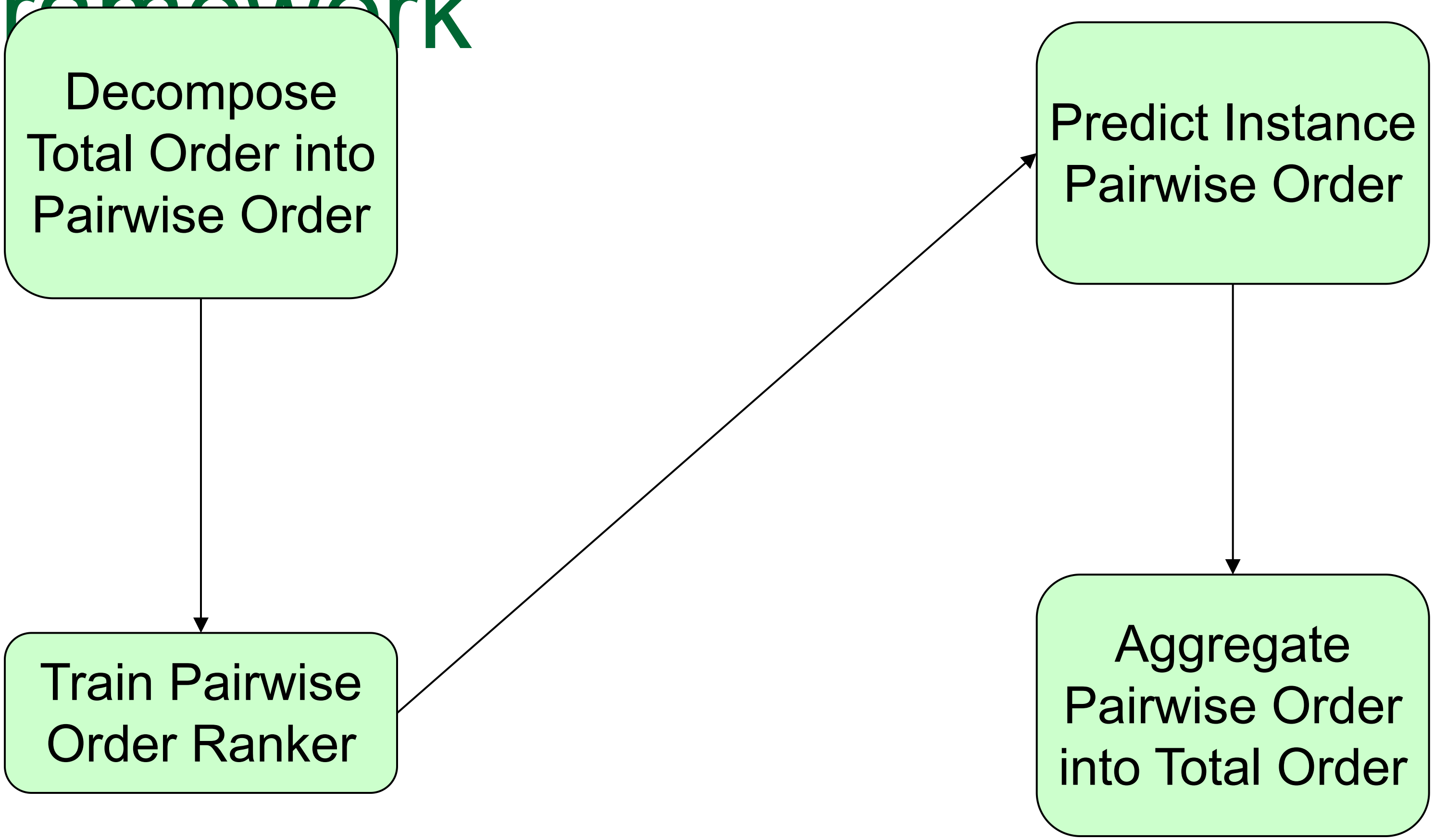
# Sample Problem

- A : Rank 3
- B : Rank 2
- C : Rank 2
- D : Rank 1

# Sample Problem

- A : Rank 3
- B : Rank 2
- C : Rank 2
- D : Rank 1
- Total Order
  - (A, B, C, D)
  - (A, C, B, D)

# General Ranking SVM Framework



# Decompose into Pairwise Order

Decompose Total Order into Pairwise Order

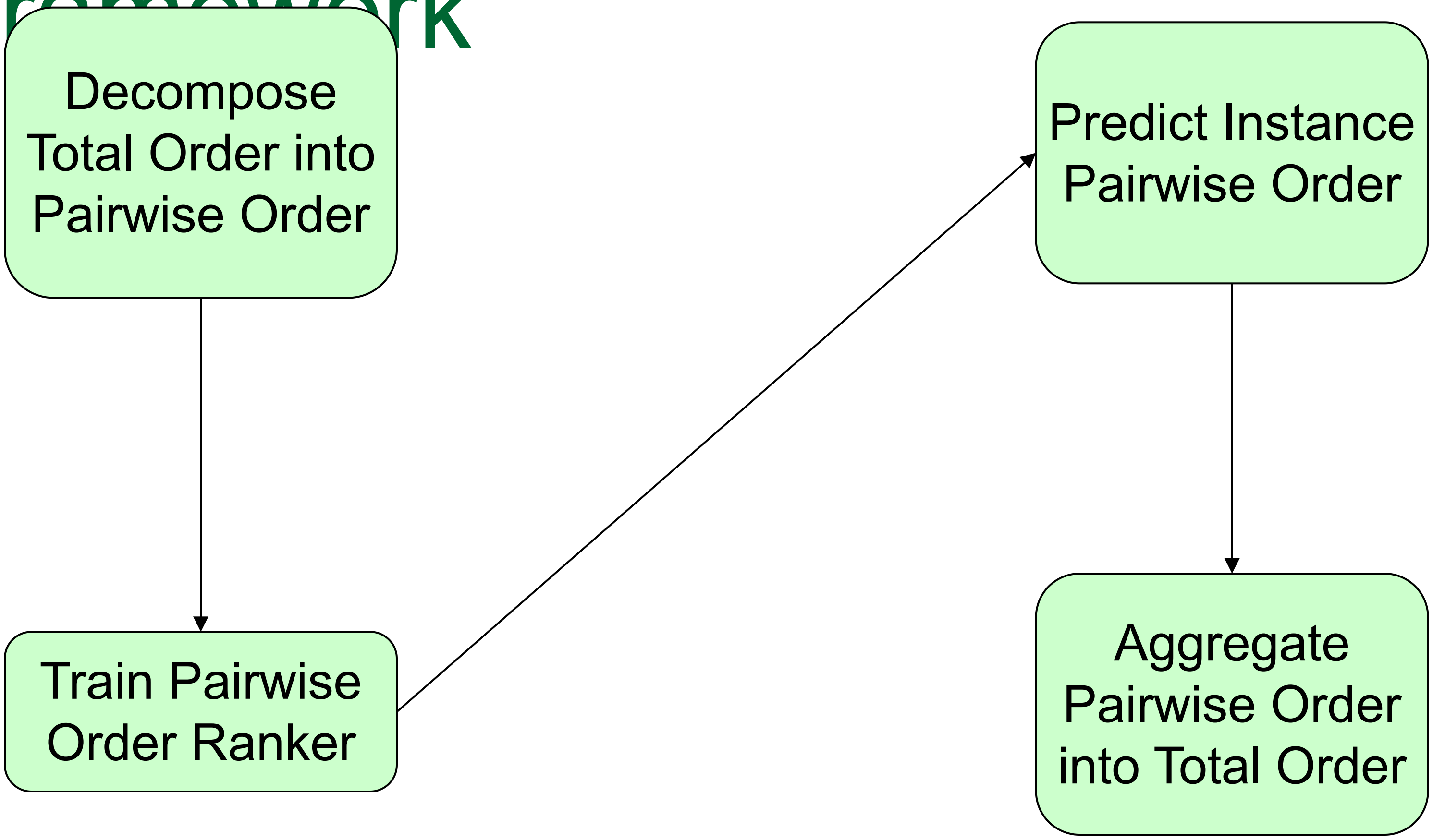
- Training Instance
  - A : Rank 3, B : Rank 2
    C : Rank 2, D : Rank 1

# Decompose into Pairwise Order

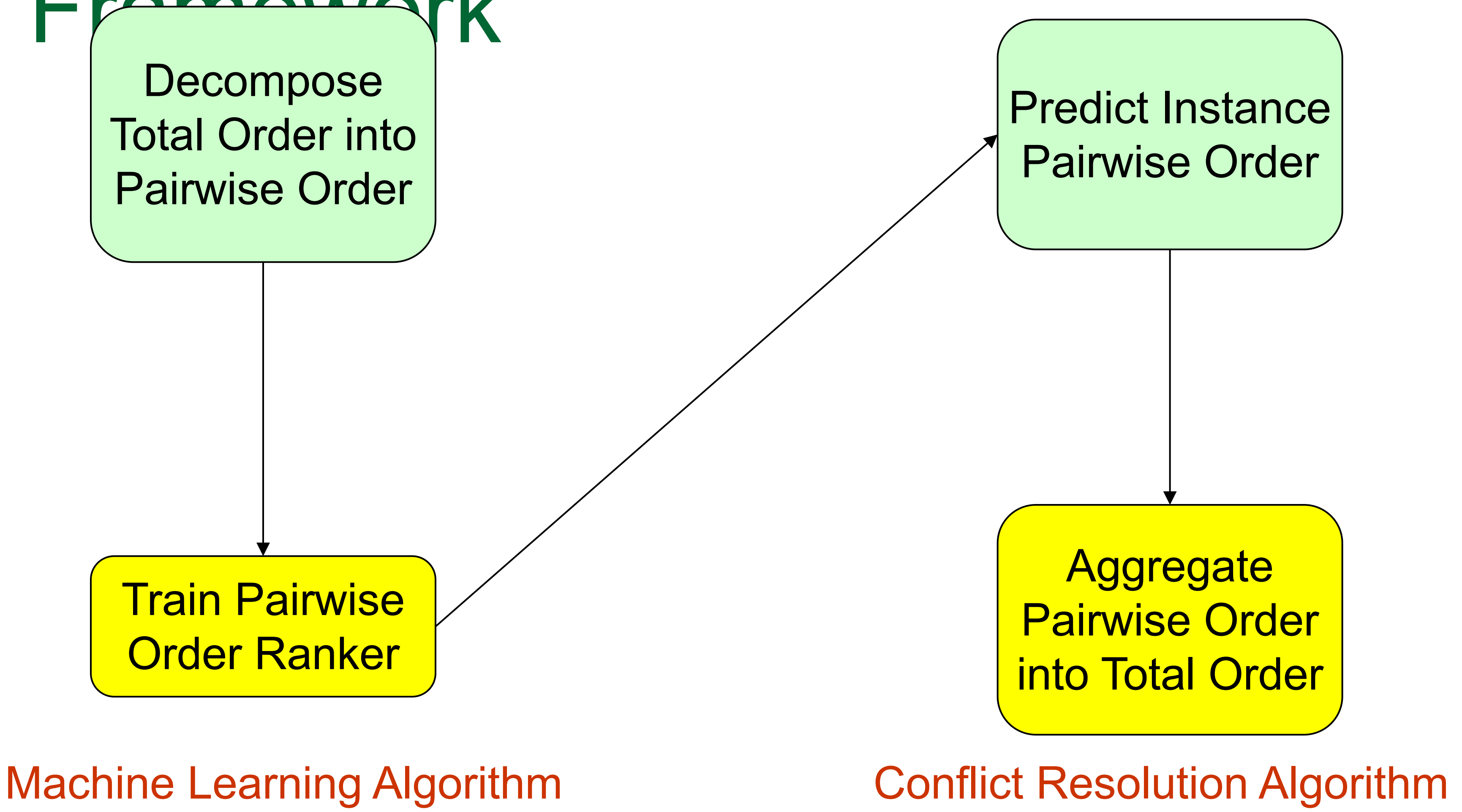
Decompose Total Order into Pairwise Order

- Training Instance
  - A : Rank 3, B : Rank 2
    C : Rank 2, D : Rank 1
- Pairwise Order
  - (A, B), (A, C), (A, D)
    (B, D), (C, D)

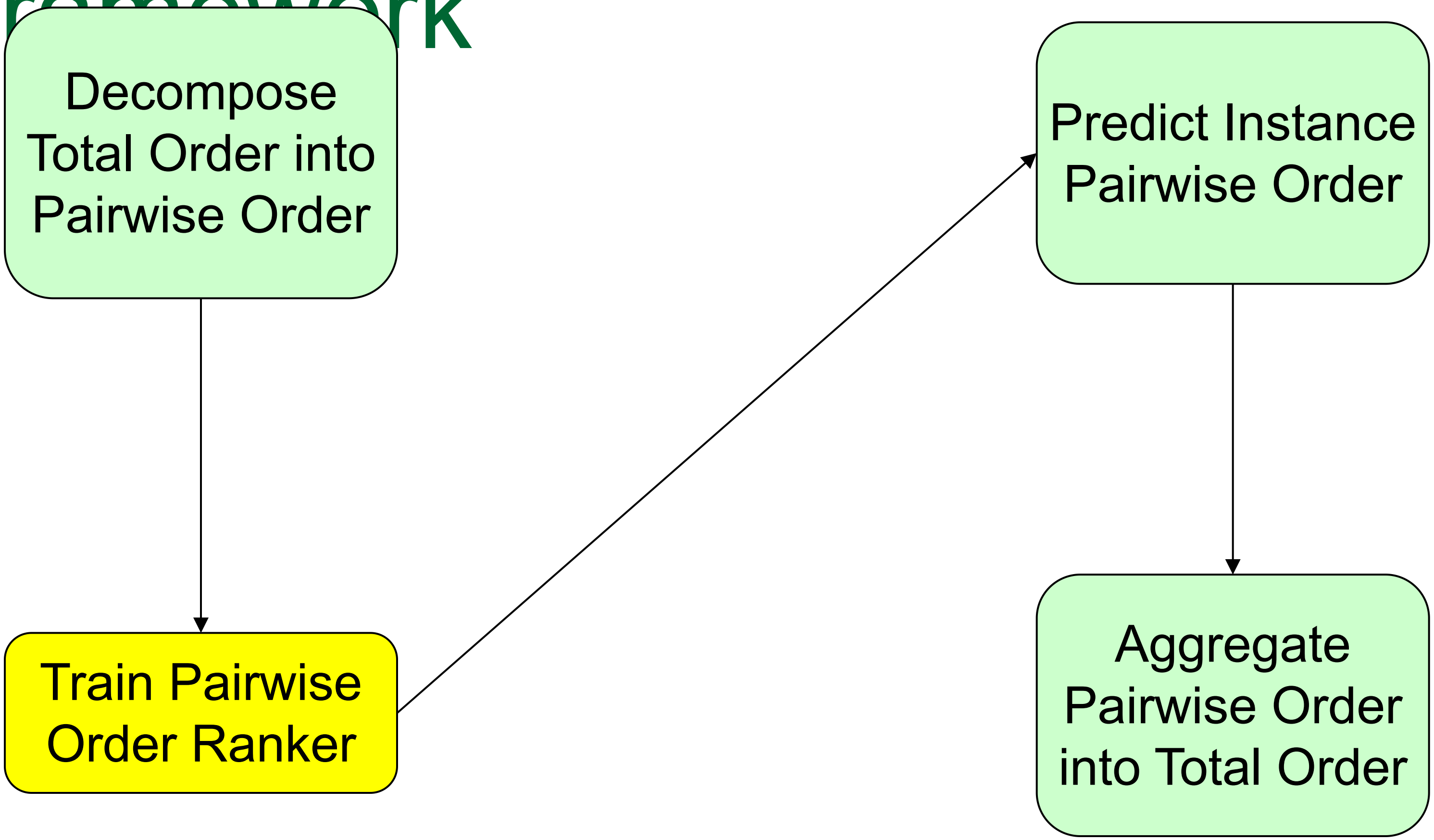
# General Ranking SVM Framework



# General Ranking SVM Framework



# General Ranking SVM Framework



# Pairwise Order Ranker

- Object X : feature vector *x*
- Object Y : feature vector *y*
- Ranker f(*x*, *y*)
  - X > Y ➔ f(*x*, *y*) > 0
  - X < Y ➔ f(x, *y*) < 0

Train Pairwise Order Ranker

# SVM Ranker

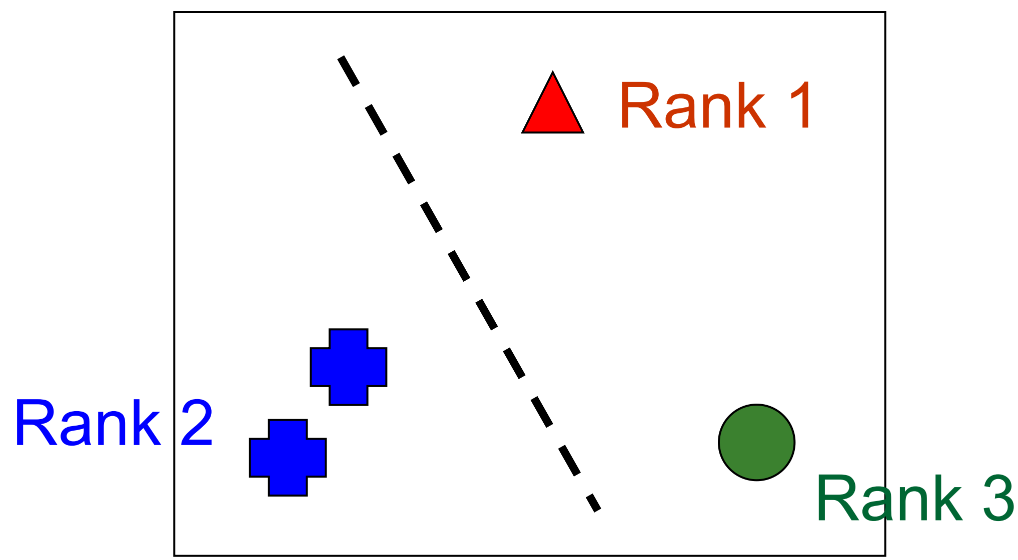
- $f(x, y) = w^T(x - y)$
- SVM finds $w$ for us

Train Pairwise Order Ranker

# Problem with Ranking SVM

- Time Complexity $O(n^k)$
- Instances not separable by a single hyperplane

Train Pairwise Order Ranker



# Solution: Multiple SVM Rankers

Train Pairwise Order Ranker

- One SVM classifier for each rank pair
- A : Rank 3, B : Rank 2

  C : Rank 2, D : Rank 1
- SVM for Rank 3 and Rank 2
  - (A, B), (A, C)
- SVM for Rank 3 and Rank 1
  - (A, D)
- SVM for Rank 2 and Rank 1
  - (B, D), (C, D)

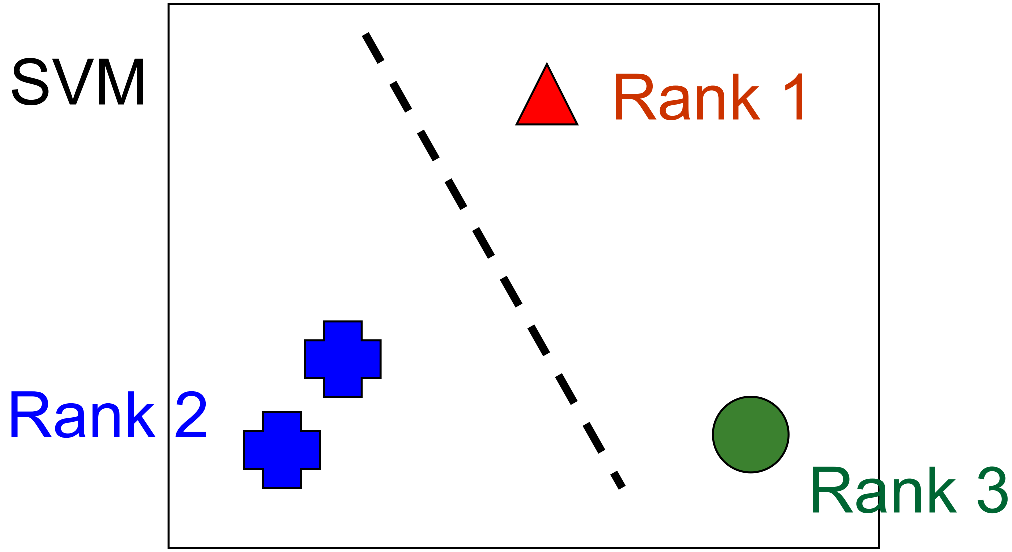
# Single VS. Multiple SVM

Train Pairwise Order Ranker

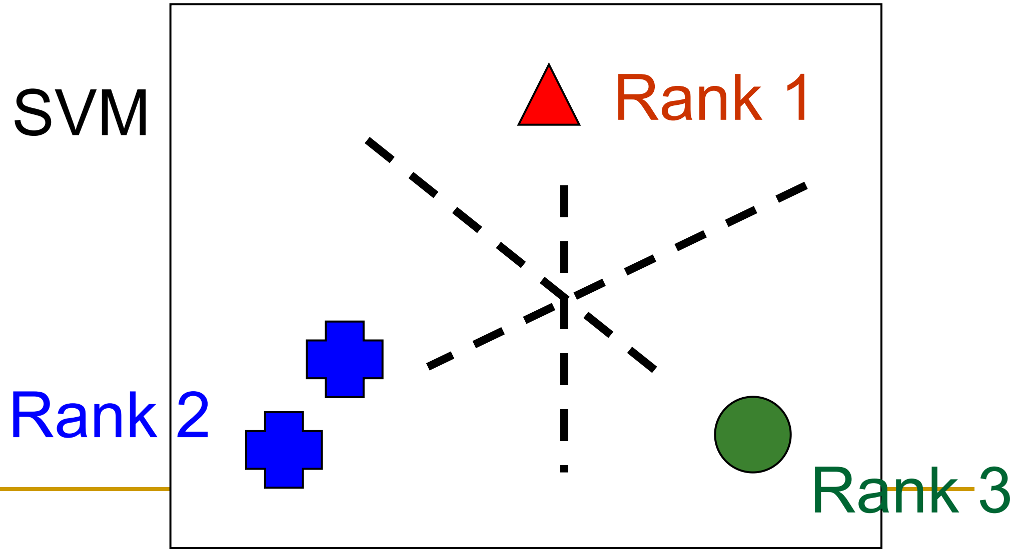
- Training Instances
  - (A, B), (A, C), (A, D) (B, D), (C, D)
- Single SVM Complexity
  - $5^k$
- Multiple SVM Compexity
  - $2^k + 1^k + 2^k$

# Single VS. Multiple SVM

One SVM

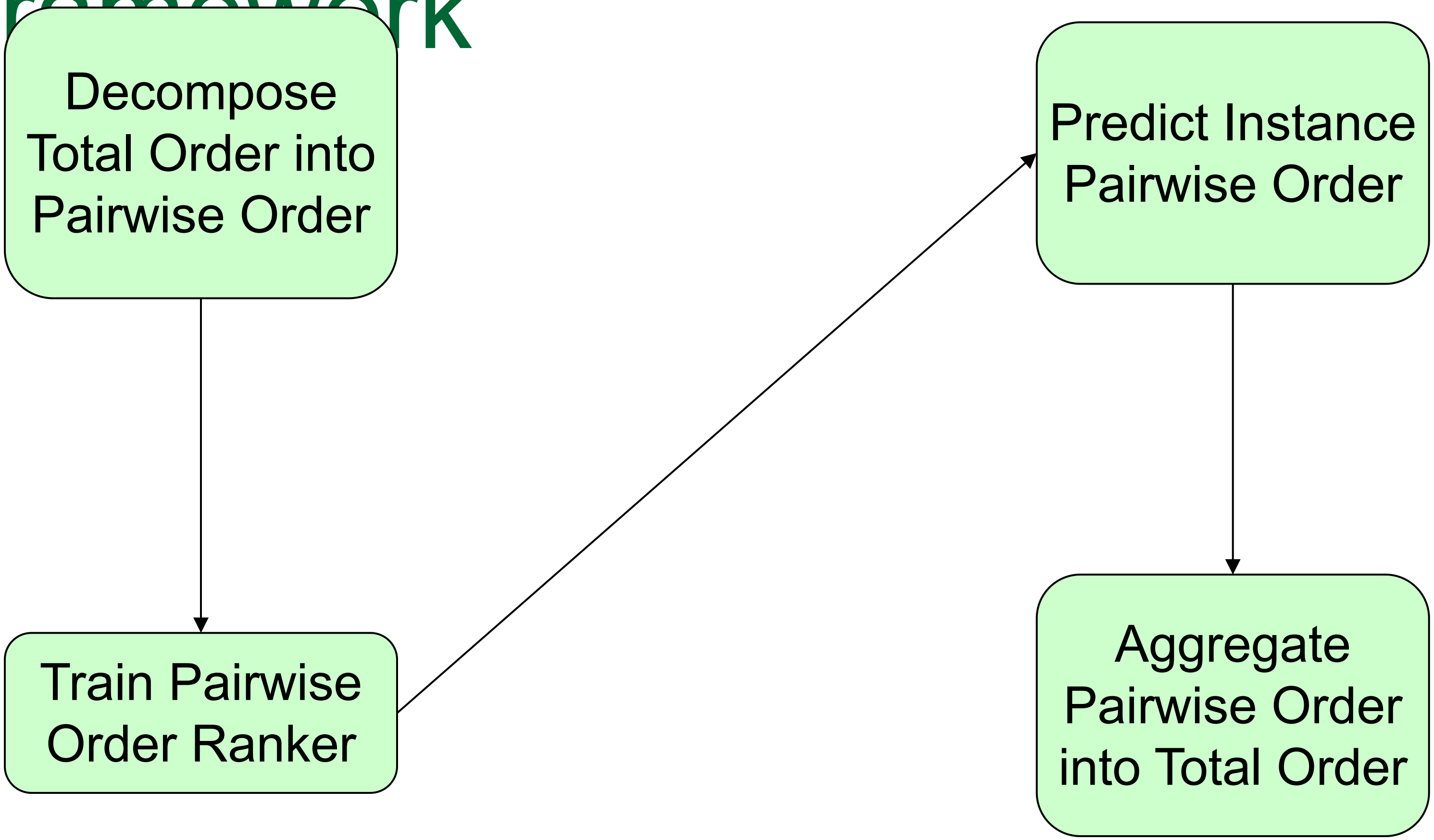



Three SVM

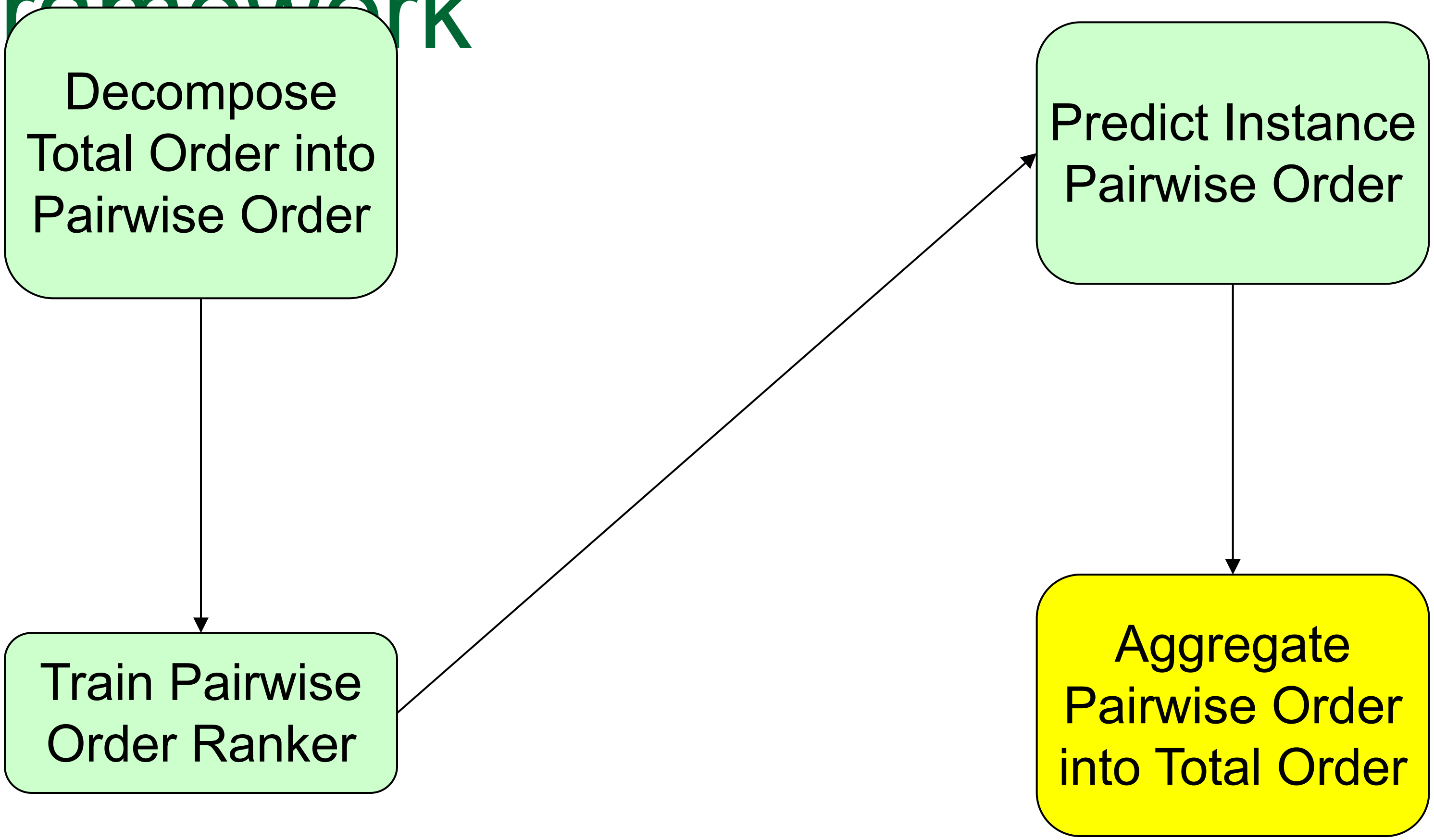



Train Pairwise Order Ranker

# General Ranking SVM Framework



# General Ranking SVM Framework



# Aggregate Pairwise Order

- Prediction Input:
  - E, F, G
- Prediction Output:
  - SVM 1: (E, F), (E, G), (F, G)
  - SVM 2: (E, F), (G, E), (F, G)
  - SVM 3: (E, F), (G, E), (G, F)
- Total Order????

Aggregate Pairwise Order into Total Order

# Weighted Borda Count

- $s(x) = \Sigma_{k=1 \text{ to } l} \alpha_k s_k(x)$
- What the???

Aggregate Pairwise Order into Total Order

# Weighted Borda Count

- $s(x) = \Sigma_{k=1 \text{ to } l}\alpha_k s_k(x)$
- What the???
- = Weighted Majority Vote

Aggregate Pairwise Order into Total Order

# Weighted Borda Count

- Prediction Output
  - SVM 1: (E, F), (E, G), (F, G)
  - SVM 2: (E, F), (G, E), (F, G)
  - SVM 3: (E, F), (G, E), (G, F)

Aggregate Pairwise Order into Total Order

# Weighted Borda Count

- Prediction Output
  - SVM 1: (E, F), (E, G), (F, G)
  - SVM 2: (E, F), (G, E), (F, G)
  - SVM 3: (E, F), (G, E), (G, F)
- E appears first 4 times

Aggregate Pairwise Order into Total Order

# Weighted Borda Count

- Prediction Output
  - SVM 1: (E, F), (E, G), (F, G)
  - SVM 2: (E, F), (G, E), (F, G)
  - SVM 3: (E, F), (G, E), (G, F)
- E appears first 4 times
- F appears first 2 times

Aggregate Pairwise Order into Total Order

# Weighted Borda Count

- Prediction Output
  - SVM 1: (E, F), (E, G), (F, G)
  - SVM 2: (E, F), (G, E), (F, G)
  - SVM 3: (E, F), (G, E), (G, F)
- E appears first 4 times
- F appears first 2 times
- G appears first 3 times

Aggregate Pairwise Order into Total Order

# Weighted Borda Count

- Prediction Output
  - SVM 1: (E, F), (E, G), (F, G)
  - SVM 2: (E, F), (G, E), (F, G)
  - SVM 3: (E, F), (G, E), (G, F)
- E appears first 4 times
- F appears first 2 times
- G appears first 3 times
- Total Order: (E, G, F)

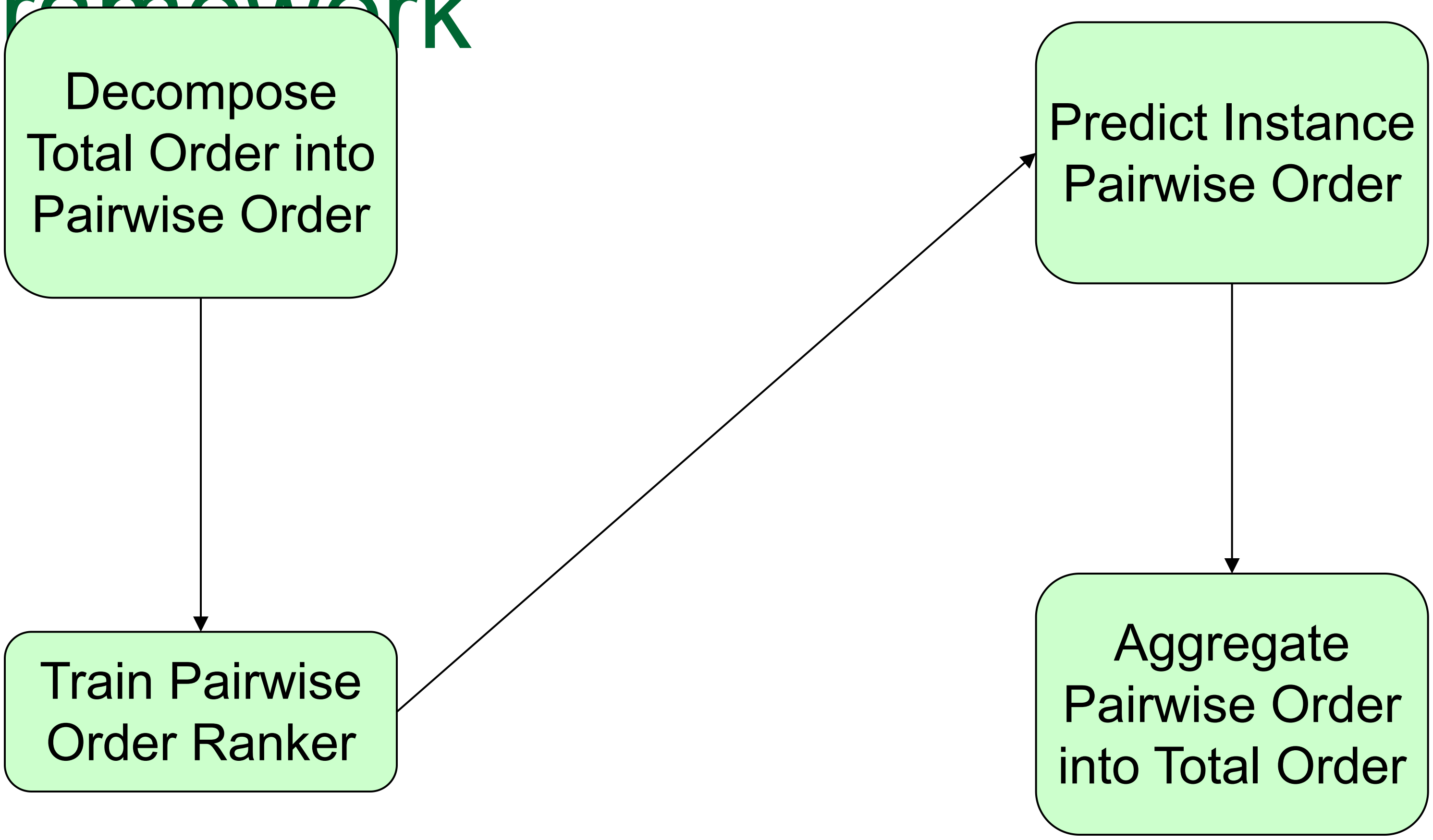
Aggregate Pairwise Order into Total Order

# Weighted Borda Count

- "Weighted" Borda Count
- Give a weight to reach SVM
- SVM 1's vote is twice as important as SVM 2's vote

Aggregate Pairwise Order into Total Order

# General Ranking SVM Framework



# Application to IR?

- Document ➔ Object
- Relevance ➔ Rank
  - Highly Relevant : Rank 3
  - Possibly Relevant : Rank 2
  - Not Relevant : Rank 1

# Corpus for Experiment

- OHSUMED
  - 348,566 documents
  - 106 queries
  - 16,140 query-document pairs
  - 3 relevance ranks

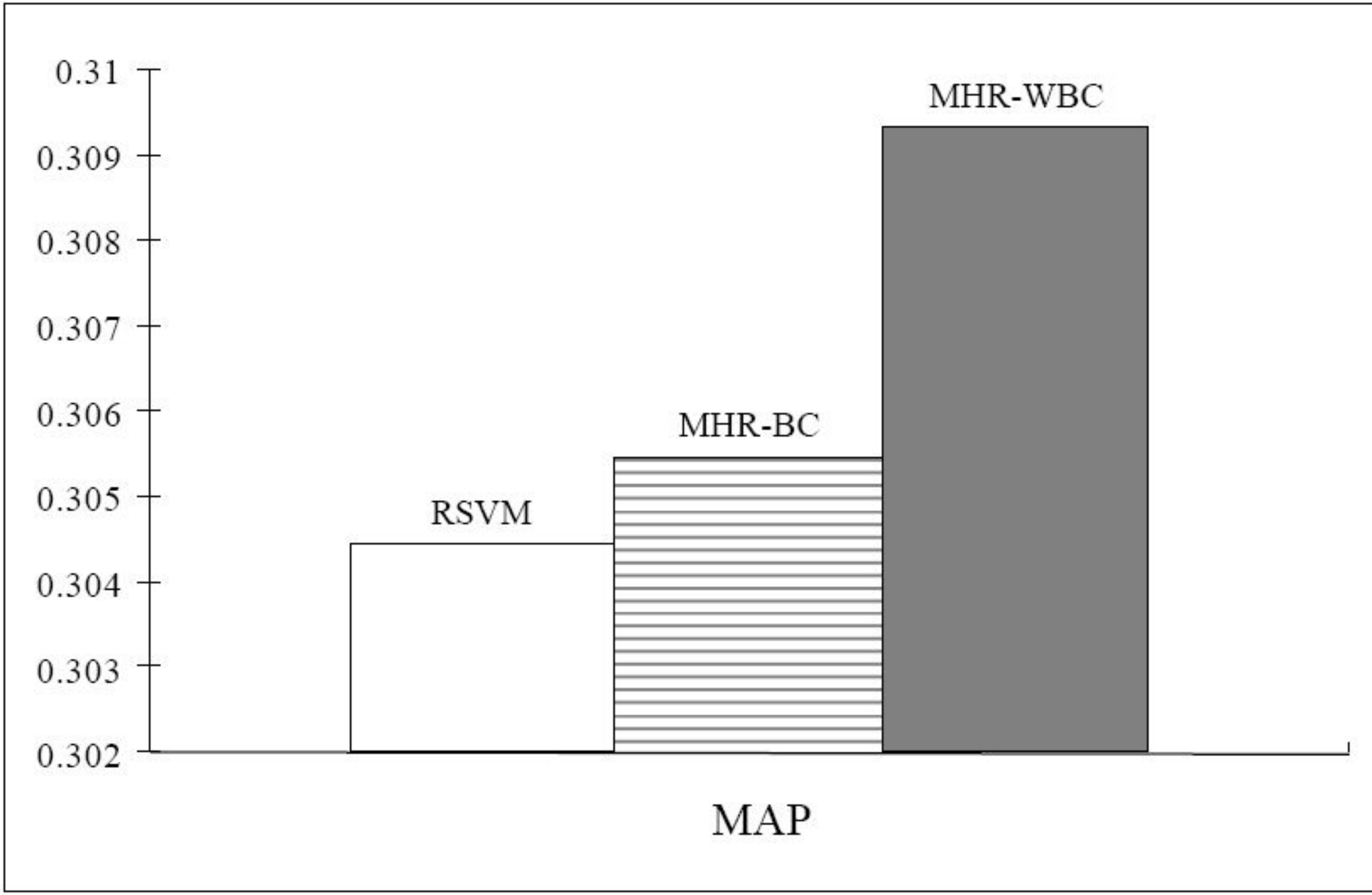
| Features | | | |
|---|---|---|---|
| 1 | $\sum_{q_i \in q \cap d} \log(c(q_i, d) + 1)$ | 2 | $\sum_{q_i \in q \cap d} \log\left(\frac{\lvert C \rvert}{c(q_i, C)} + 1\right)$ |
| 3 | $\sum_{q_i \in q \cap d} \log\left(\frac{c(q_i, d)}{\lvert d \rvert} idf(q_i) + 1\right)$ | 4 | $\sum_{q_i \in q \cap d} \log\left(\frac{c(q_i, d)}{\lvert d \rvert} + 1\right)$ |
| 5 | $\sum_{q_i \in q \cap d} \log\left(\frac{c(q_i, d)}{\lvert d \rvert} \cdot \frac{\lvert C \rvert}{c(q_i, C)} + 1\right)$ | 6 | $\sum_{q_i \in q \cap d} \log(idf(q_i))$ |
| 7 | log(BM25 score) | | |

# Corpus for Experiment

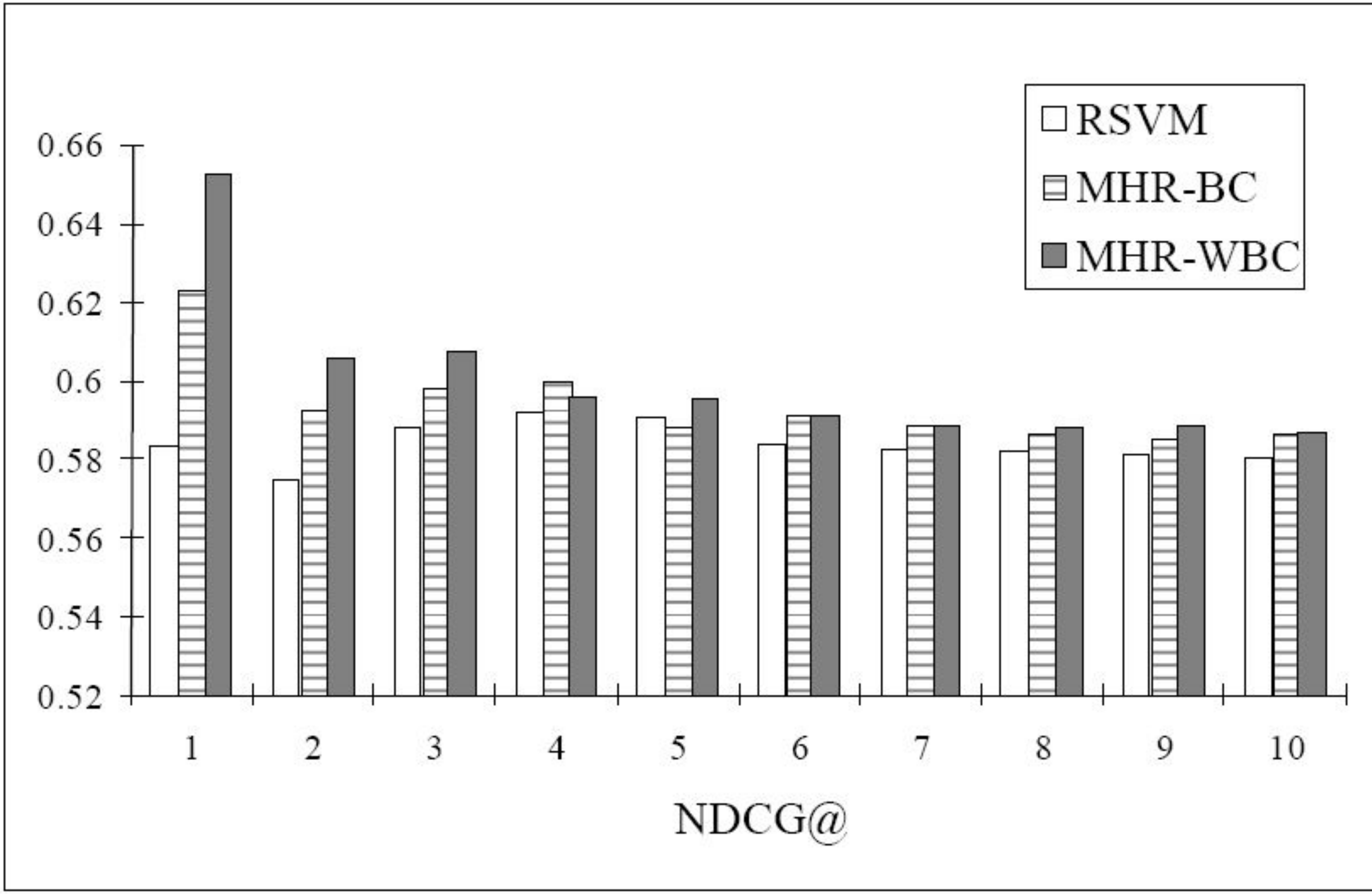
- Definition Search
  - 170 queries
  - 2,000 documents per query
  - 3 relevance ranks

1. <query> occurs at beginning of paragraph.
2. <query> begins with ‘the’, ‘a’, or ‘an’.
3. All the words in <query> begin with uppercase letters.
4. Paragraph contains predefined negative words, e.g. ‘he’, ‘said’, ‘she’
5. <query> contains pronouns.

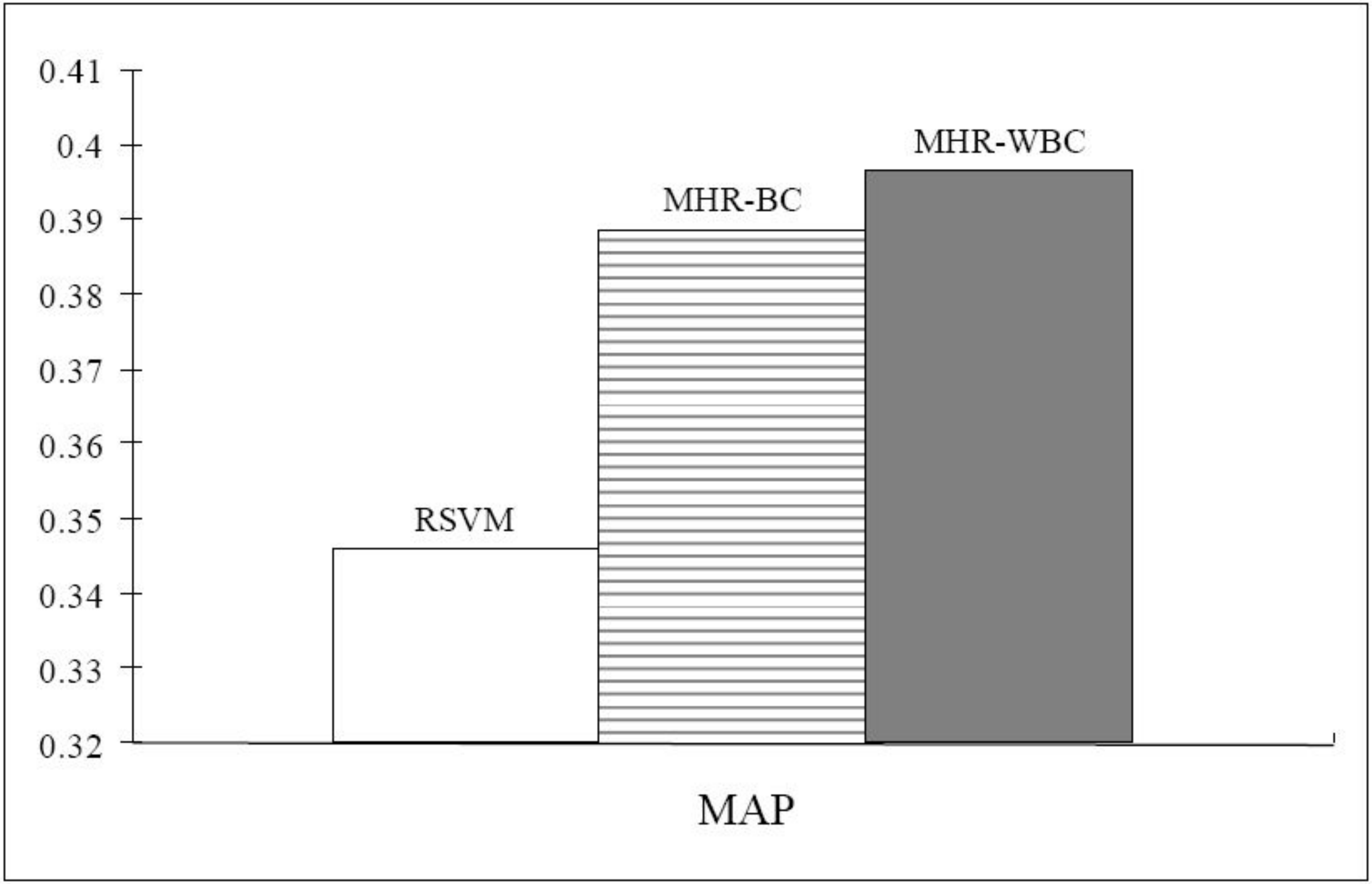
# OHSUMED Results



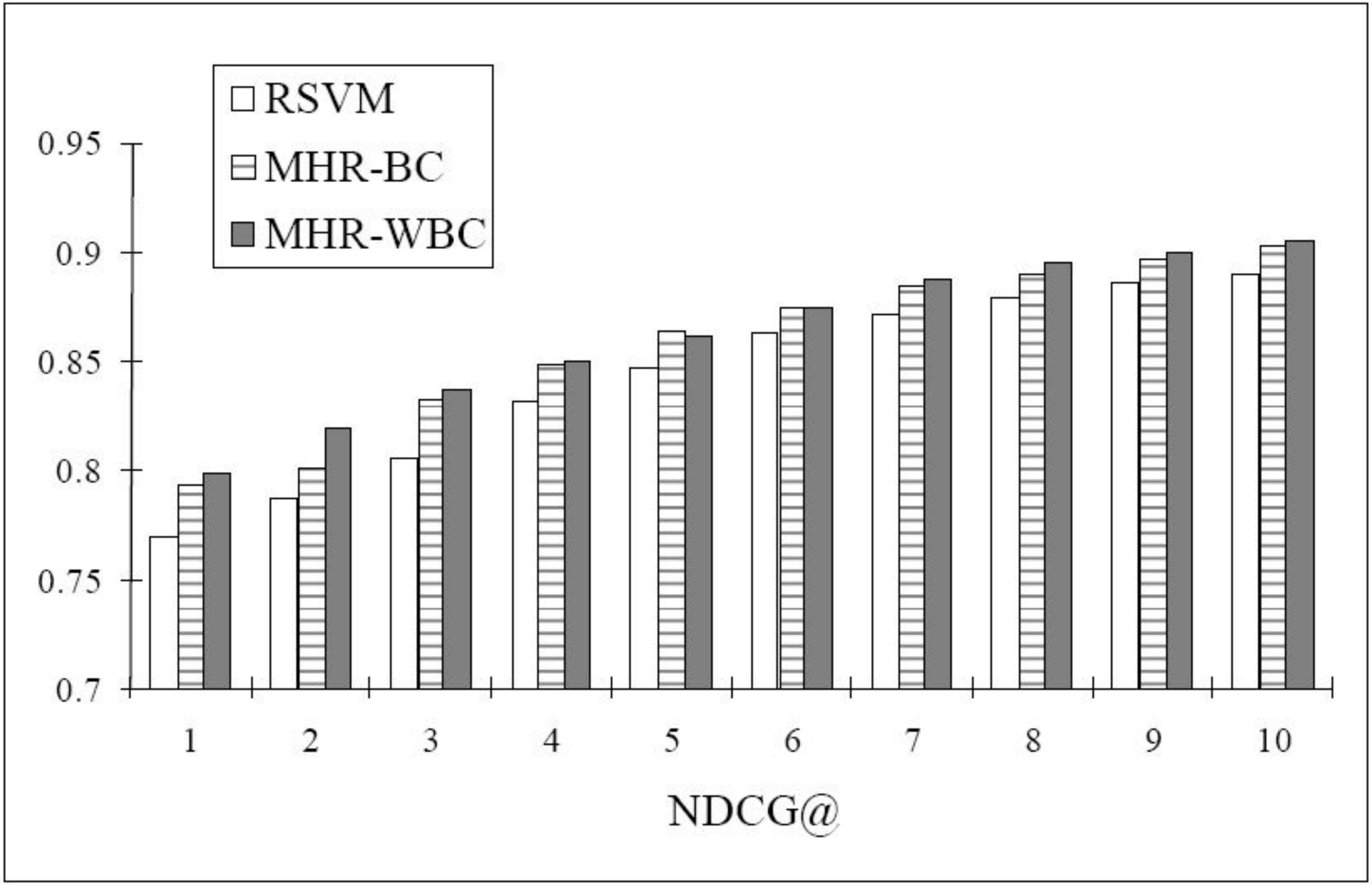
# OHSUMED Results



# Definition Search Results



# Definition Search Results



# OHSUMED Training Time

| Minutes | MHR | | | | RSVM |
|---|---|---|---|---|---|
| | $\omega_{1,2}$ | $\omega_{1,3}$ | $\omega_{2,3}$ | Sum | |
| trial 1 | 17 | 90 | 175 | 282 | 823 |
| trial 2 | 17 | 75 | 200 | 292 | 841 |
| trial 3 | 16 | 78 | 154 | 248 | 663 |
| trial 4 | 22 | 92 | 196 | 310 | 887 |

# Definition Search Training Time

| Seconds | MHR | | | | RSVM |
|---|---|---|---|---|---|
| | $\omega_{1,2}$ | $\omega_{1,3}$ | $\omega_{2,3}$ | Sum | |
| trial 1 | 0.07 | 0.06 | 1.17 | 1.30 | 1.90 |
| trial 2 | 0.11 | 0.07 | 2.78 | 2.96 | 3.10 |
| trial 3 | 0.13 | 0.07 | 1.41 | 1.61 | 2.83 |
| trial 4 | 0.08 | 0.07 | 1.67 | 1.82 | 3.76 |

# Contribution

- Multiple Hyperplanes are good because
  - Less training time
  - More accurate

# General Ranking SVM Framework