

Progressive Comparison for Ranking Estimation (IJCAI-16)

Ryusuke Takahama Kyoto University / JST, ERATO

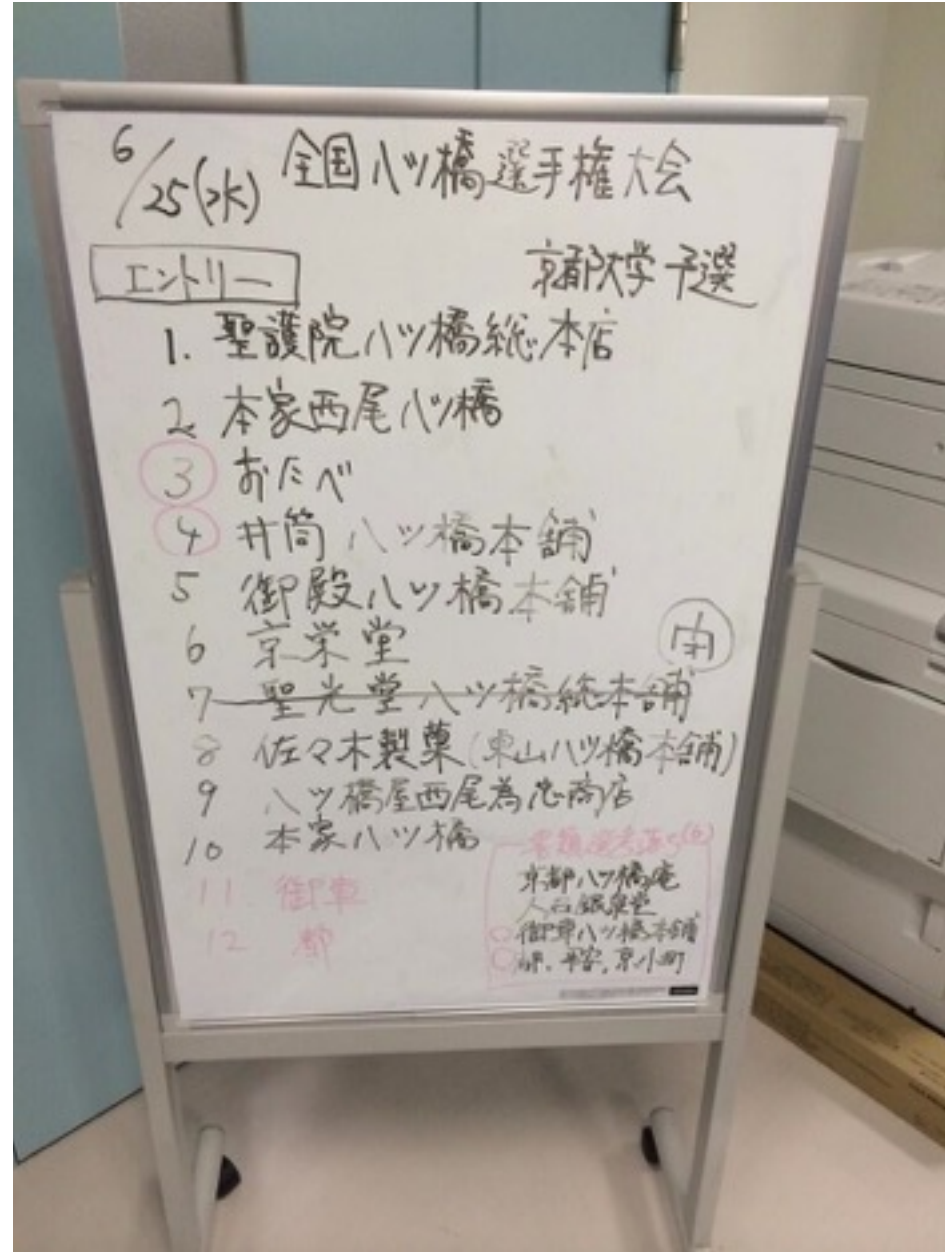
Toshihiro Kamishima AIST

Hisashi Kashima Kyoto University

Outline

- Background
- Problem setting
- **Progressive Comparison**
- **Active learning** of Progressive Comparison
 - Change in Distributions
 - Change in Winning Probabilities
- Experimental results

The 1st All Japan Yatsuhashi Championship



私たちはお土産にどの八ッ橋を買えばよいのか

Which is the Best Yatsuhashi Brand for a Souvenir?

高濱 隆輔 * ¹	大谷 直樹 * ¹	横井 祥 * ¹	荒井 智裕 * ¹	則 のぞみ * ²	鵜飼 紀衣 * ²
Ryusuke Takahama	Naoki Otani	Sho Yokoi	Tomohiro Arai	Nozomi Nori	Norie Ugai
		中澤 巧爾 * ²	鹿島 久嗣 * ²		
		Koji Nakazawa	Hisashi Kashima		

*¹ 京都大学工学部情報学科 計算機科学コース ソフトウェア基礎論分野
Group of Foundation of Software Science, Undergraduate Course Program of Computer Science
Undergraduate School of Informatics and Mathematical Science, Faculty of Engineering, Kyoto University

*² 京都大学大学院情報学研究科 知能情報学専攻 ソフトウェア基礎論分野
Group of Foundation of Software Science, Department of Intelligence Science and Technology
Graduate School of Informatics, Kyoto University



れを買っても同じなのかという問いに答えるべく、数理モデルを用いた検証を行った。具体的には、先行研究 [e 京都 10] に基づき、京都市内で販売されている八ッ橋のうち主要なもの 11 銘柄を、9 人の評価者が実際に試食、評価を行い、これを

- (1) 勝率
- (2) 主固有ベクトル [Keener 93]
- (3) Bradley-Terry モデル [Bradley 52]

地域 > 近畿

2015年3月17日

京都大学の研究成果論文「私たちはお土産にどの八ッ橋を買えばよいのか」

多くの人の悩みが解決される...かもしれない。


ハッ橋 ハつ橋 土産 京都 京都大学 論文 研究 研究成果論文

somali_bleu 118641view 31コメント 2344 B! 526 f 6499 お気に入り



301


まとめ メニュー

 Eriko Amano @sabarya 2015-03-16 22:51:46

京都大学情報学研究科・工学部情報学科の研究成果です。(PDF) / “私たちはお土産にどの八ッ橋を買えばよいのか” htn.to/4JSMUx

 うさこじぞう @usako_jizou 2015-03-17 08:04:01

「私たちはお土産にどの八ッ橋を買えばよいのか」
京大工学部の論文でしたよ。しっかりとした実験に基づいた論文だ....
[ml.ist.i.kyoto-u.ac.jp/wp/wp-content/...](http://ml.ist.i.kyoto-u.ac.jp/wp/wp-content/)

 数土直志 @sudotadashi 2015-03-17 09:19:21

長年の悩みが解決する(^.^) @sabarya: 京都大学情報学研究科・工学部情報学科の研究成果です。(PDF) / “私たちはお土産にどの八ッ橋を買えばよいのか” htn.to/4JSMUx

トゥギャトピ! 20時03分更新
編集部がオススメまとめをピックアップしてお届け中



旬名人に聞いた「本当の旬の食べ方」
生でも食べれる?スーパーの旬とは違う真の旬とは

井の頭公園にいるブルースじいさんが渋いNEW!
真のブルースマンによる圧巻の演奏が話題に!

#世界猫の日 でさまざまな猫が大集合!NEW!
かわいい猫、美人な猫...橋本環奈や阿部サダヲまで!?


8/8にバインアメで全力で遊ぶ各公式アカNEW!
コラ画像を作ったりチャーハンに混ぜたり

関西のいろんな日本酒を京都で飲み比べ!NEW!
お酒だけではなく京野菜料理も目白押しのフェスです!


熊本城や神社仏閣の被害状況が見える旅路NEW!
観光を楽しみながら、復興を祈ります

リアルタイムランキング 7分40秒前更新

1 **「天皇陛下お気持ち表明」** フォントは放送開始14分後、皿は27分後にほぼ特定、そして難関と思われた石も



2 某「私信」回答に対する反応



地域 > 近畿

2015年3月17日

京都
えは

118642view

どこのハッ橋を買

多くの

ハッ橋 ハッ橋 土産 京 京都大学 論文 研究 研究成果論文

somali_bleu 118641view 31コメント

2344 B! 526 f 6499 お気に入り



301

まとめ

2344

526

6499



Erik
京
ち

うさこじぞう @usako_jizou

2015-03-17 08:04:01

「私たちはお土産にどのハッ橋を買えばよいのか」
京大工学部の論文でしたよ。しっかりとした実験に基づいた論文だ....
[ml.ist.i.kyoto-u.ac.jp/wp/wp-content/...](http://ml.ist.i.kyoto-u.ac.jp/wp/wp-content/)



数土直志 @sudotadashi

2015-03-17 09:19:21

長年の悩みが解決する(^.^)*@sabarya: 京都大学情報学研究科・工学部情報学科の研究成果です。(PDF) / 「私たちはお土産にどのハッ橋を買えばよいのか」 htn.to/4JSMUx

トウギャトビ! 20時03分更新
編集部がオススメまとめをピックアップしてお届け中



旬名人に聞いた「本当の旬の食べ方」
生でも食べれる?スーパーの旬とは違う真の旬とは

お酒だけではなく京野菜料理も目白押しのフェスです!
熊本城や神社仏閣の被害状況が見える旅路NEW!
観光を楽しみながら、復興を祈ります

リアルタイムランキング 7分40秒前更新

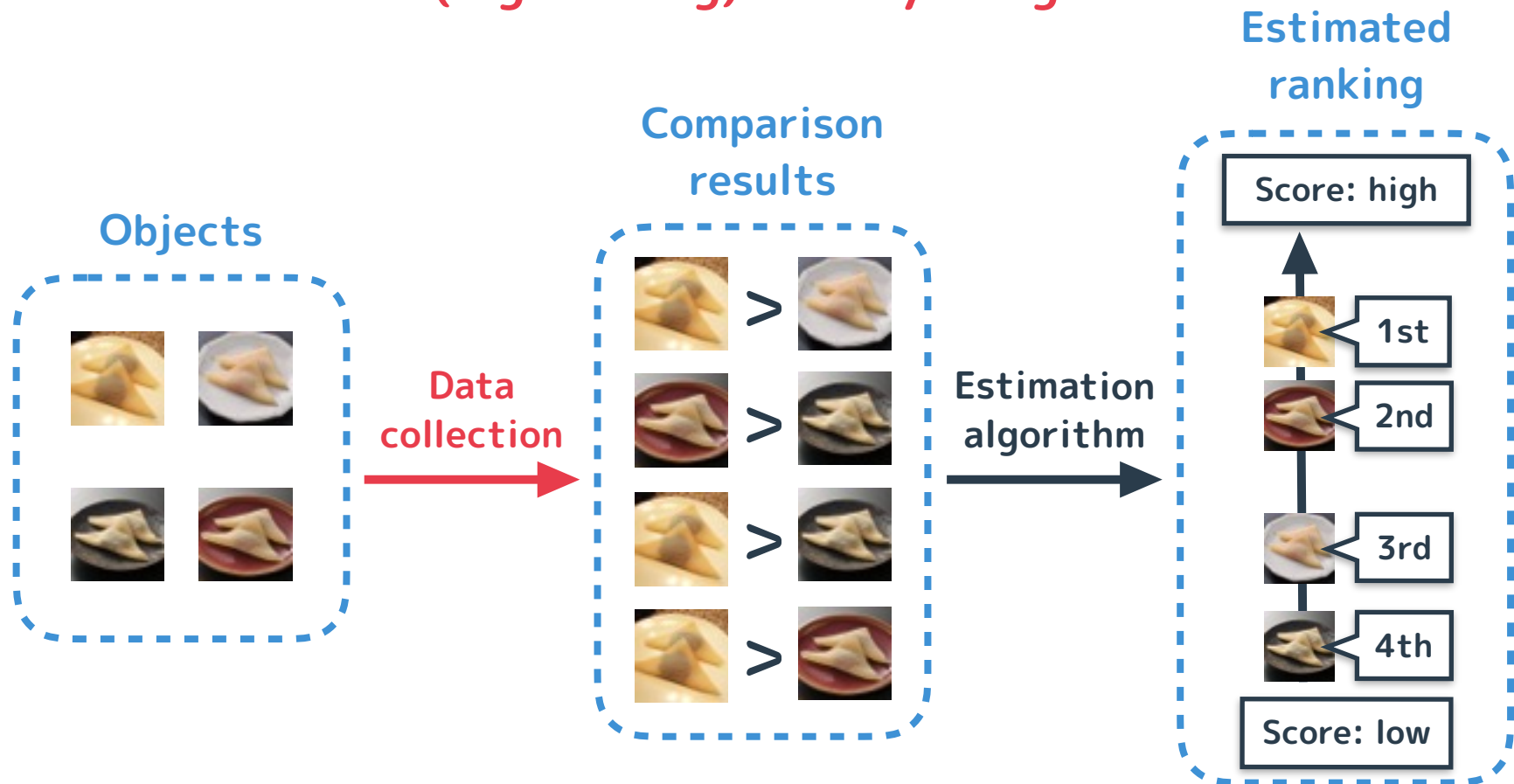
1 「天皇陛下お気持ち表明」 フォントは放送開始14分後、皿は27分後にほぼ特定、そして難関と思われた石も

2 某「私信」回答に対する反応

Background:

The 1st All Japan Yatsushashi Championship

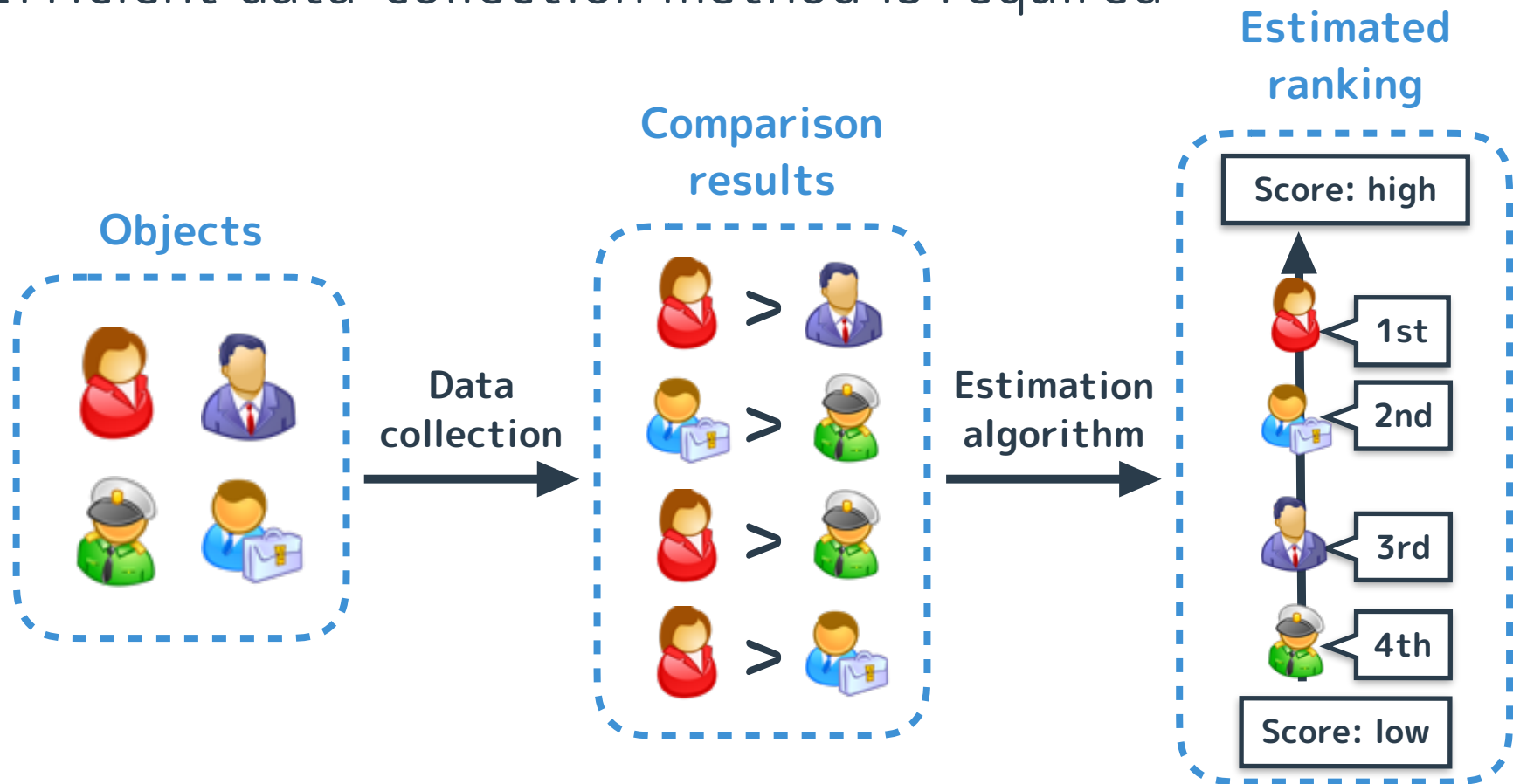
- Estimate Yatsushashi ranking using pairwise comparison data
 - **Data-collection (e.g. eating) is very tough**



Problem setting:

Estimate ranking from pairwise comparison data

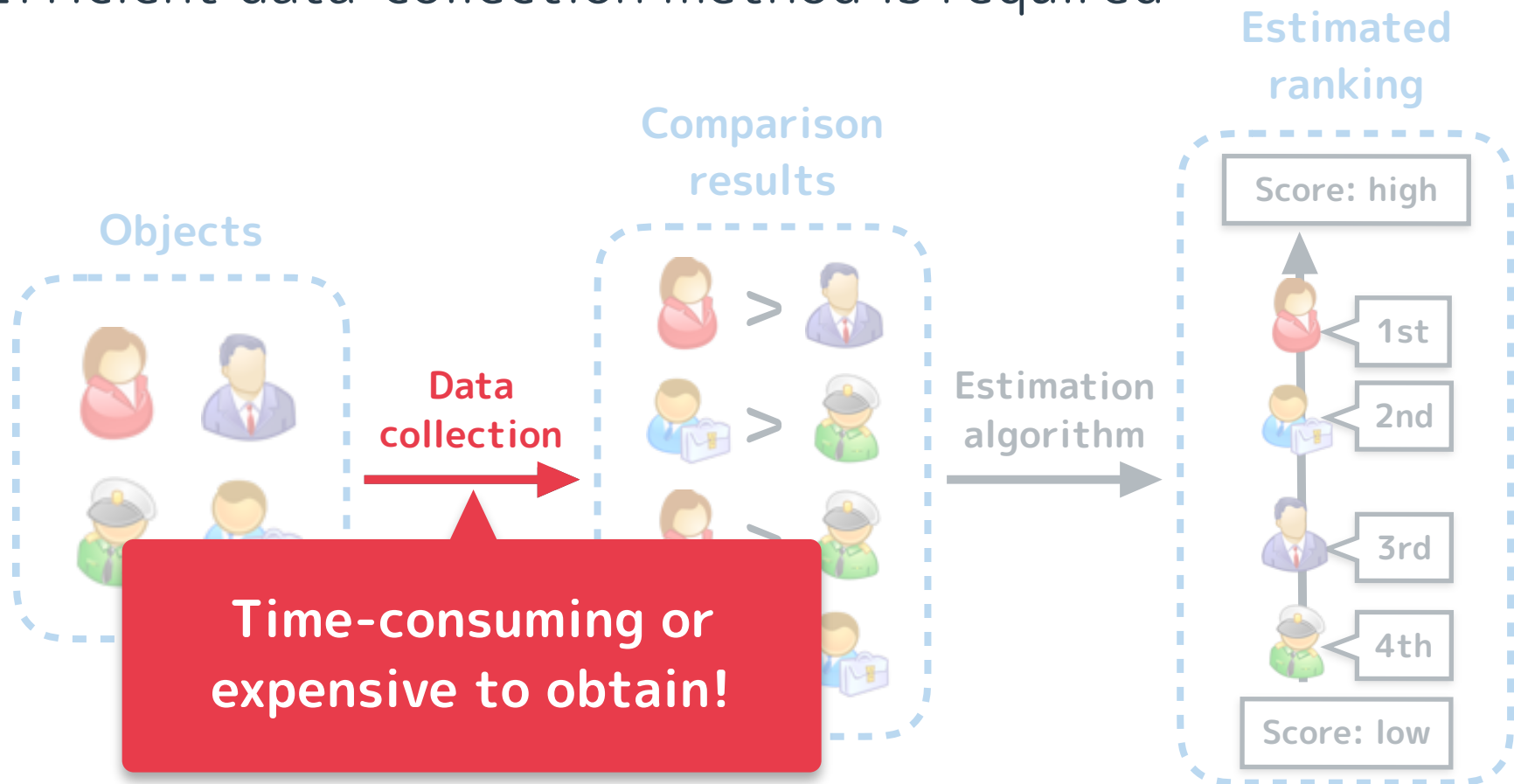
- Estimate objects ranking using pairwise comparison data
- Efficient data-collection method is required



Problem setting:

Estimate ranking from pairwise comparison data

- Estimate objects ranking using pairwise comparison data
- Efficient data-collection method is required



Problem setting:

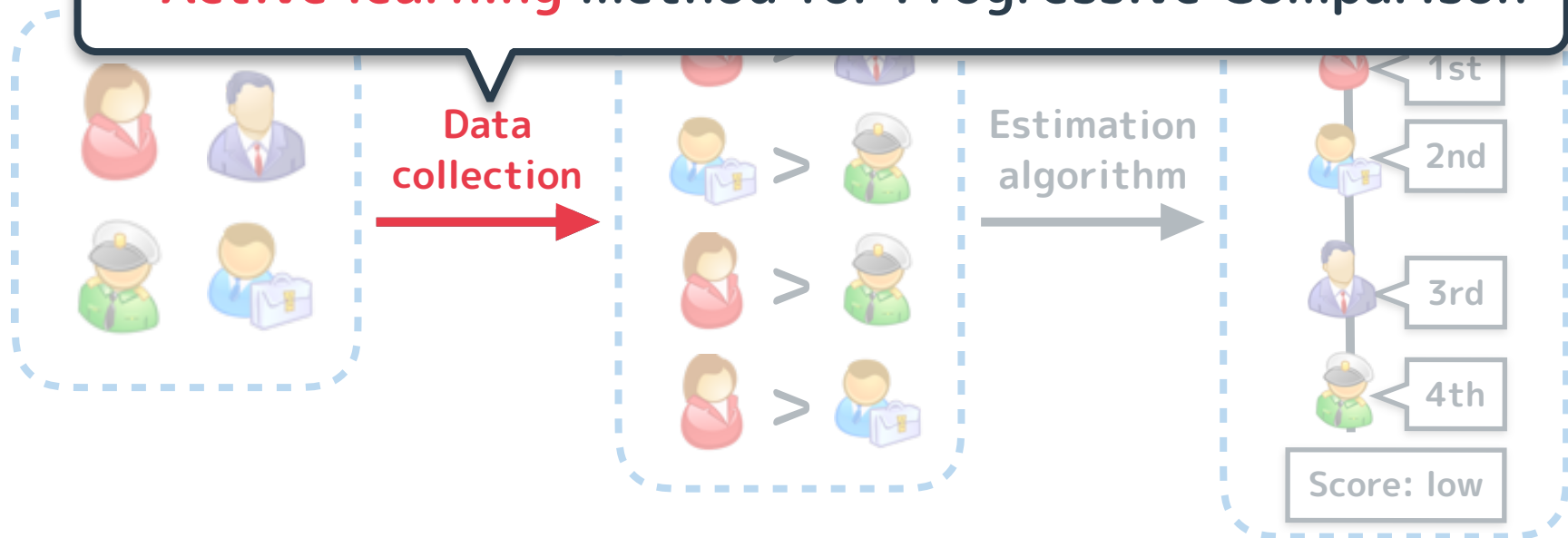
Estimate ranking from pairwise comparison data

- Estimate objects ranking using pairwise comparison data
- Efficient data-collection method is required

Estimated

Proposed to collect data efficiently:

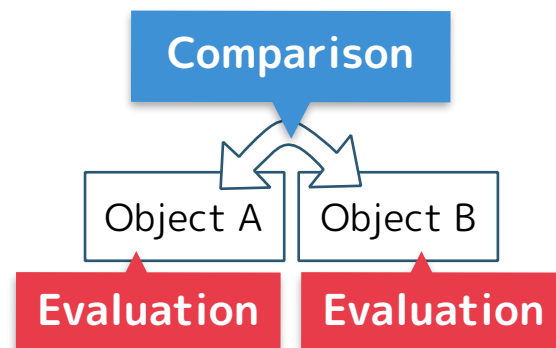
- **Progressive Comparison**
- **Active learning** method for Progressive Comparison



Model of evaluators:

Evaluators evaluate objects and then compare them

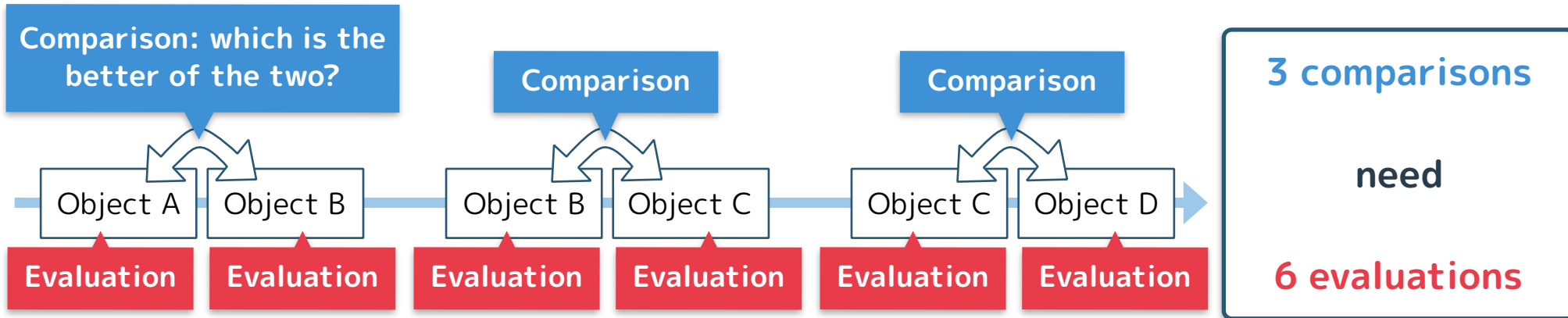
- Definition:
 - **Evaluation:** giving an (internal) score to an object
 - **Comparison:** determining the winner based on the two objects scores to create one comparison result
- Assumption:
 - **The cost of an evaluation is substantially larger** than that of a comparison
 - **Evaluators can not remember many internal scores** of objects



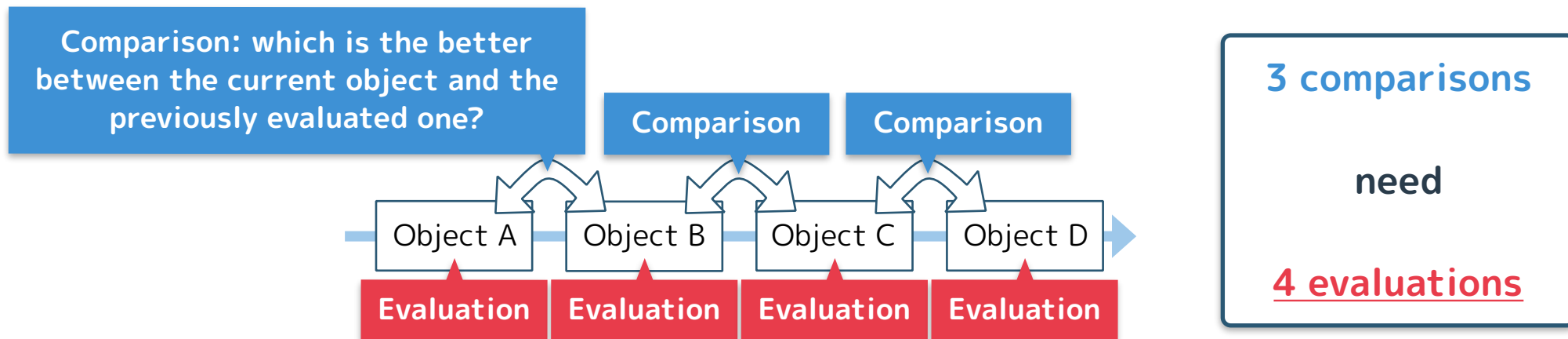
Progressive Comparison:

Data-collection method needing fewer evaluations

- Existing method (Standard pairwise comparison):



- Proposed method (**Progressive Comparison**):

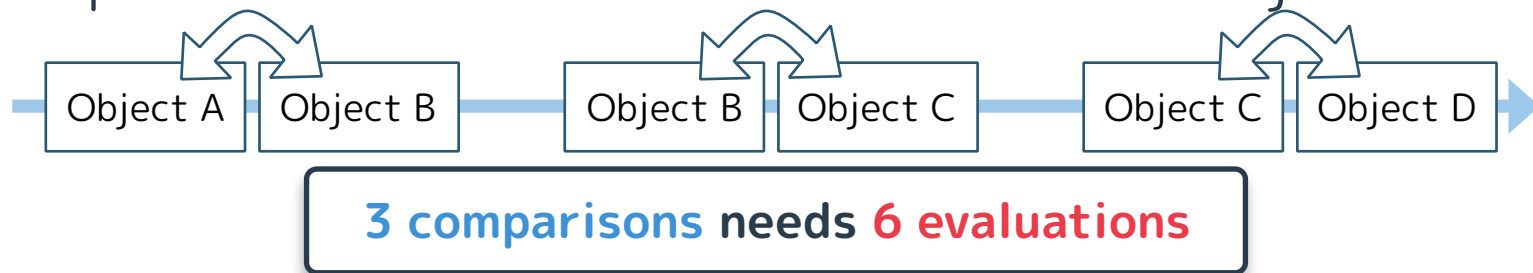


Standard pairwise v.s. Progressive:

Progressive Comparison needs fewer evaluations

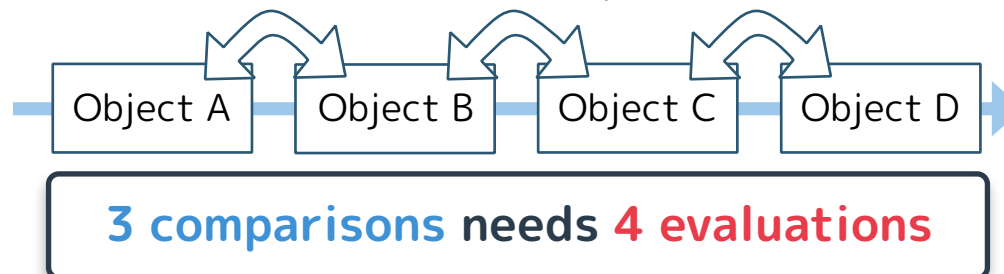
- Standard pairwise comparison:

- N comparison results need **2N evaluations** of objects



- Progressive Comparison:

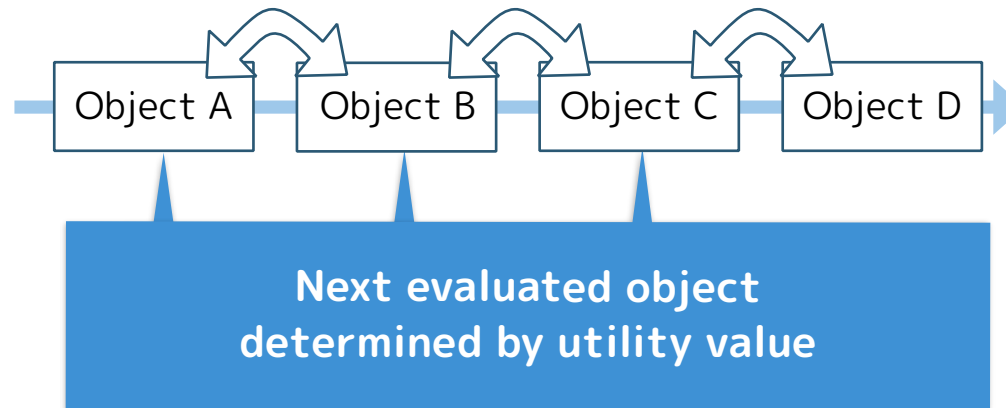
- N comparison results need **(N + 1) evaluations** of objects
- However, this method has a constraint: each object must be compared in two consecutive comparisons



Active learning for Progressive comparison:

Estimate ranking efficiently by selecting pairs

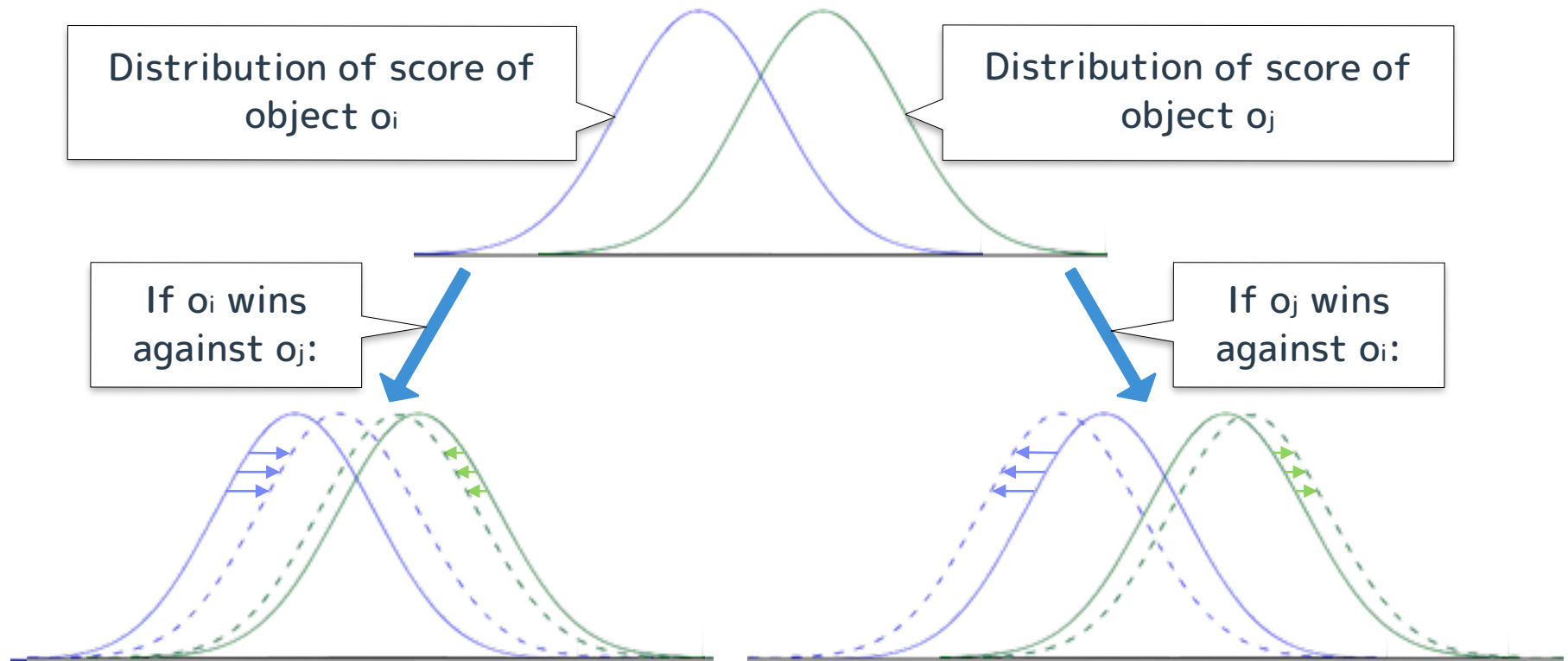
- **Utilities** calculated for each pair
- Priority given to a pair that has larger utility value
- Two definitions of utility proposed:
 - Change in Distributions (CiD):**
expectation of changes in distributions of object scores
 - Change in Winning Probabilities (CiWP):**
expectation of changes in winning probability matrices



(i) Change in Distributions (CiD):

Calculate expectation of changes in distributions

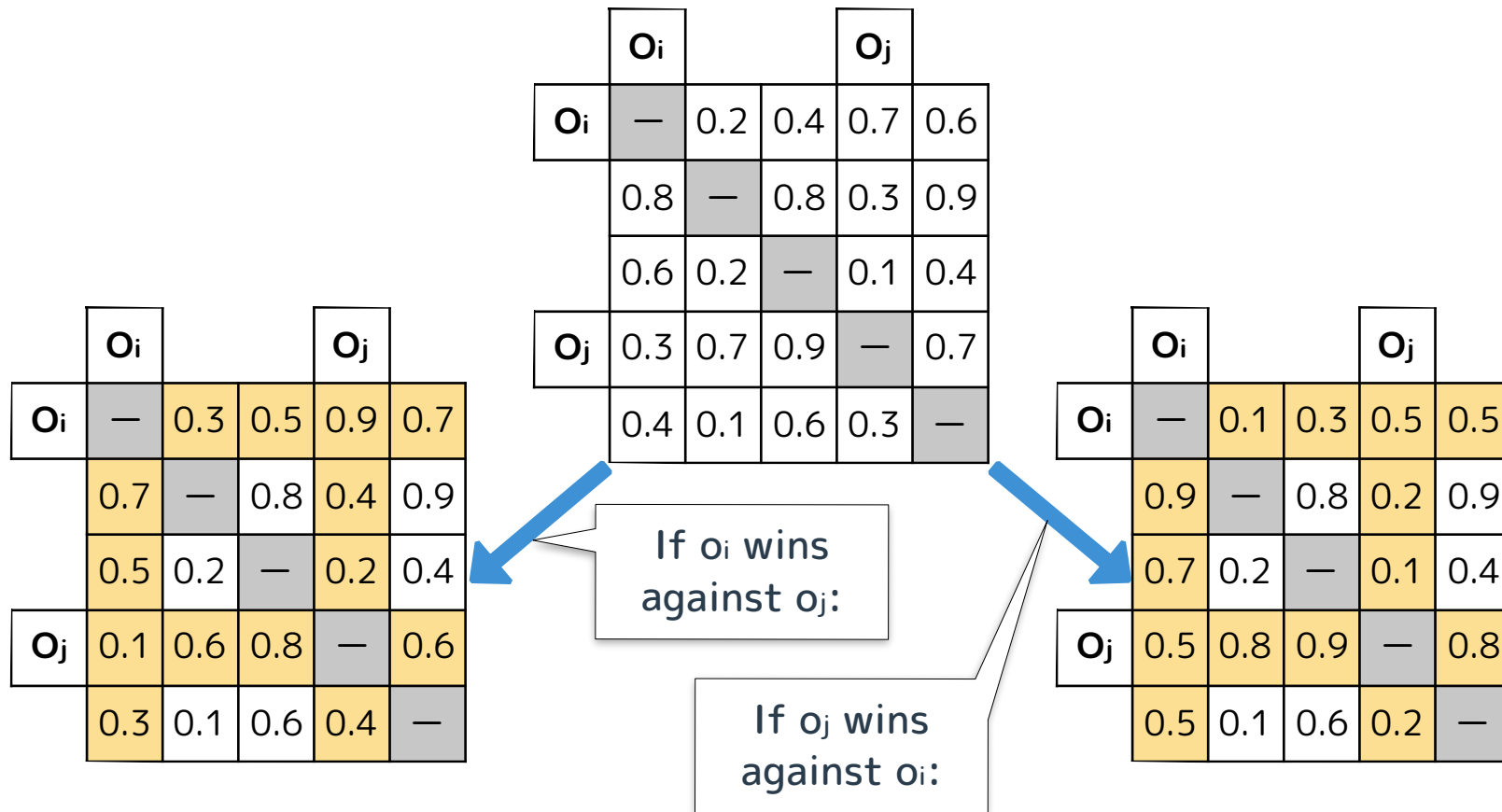
- Expectation of changes in distributions calculated by KL divergence between normal distributions:



(ii) Change in Winning Probabilities (CiWP):

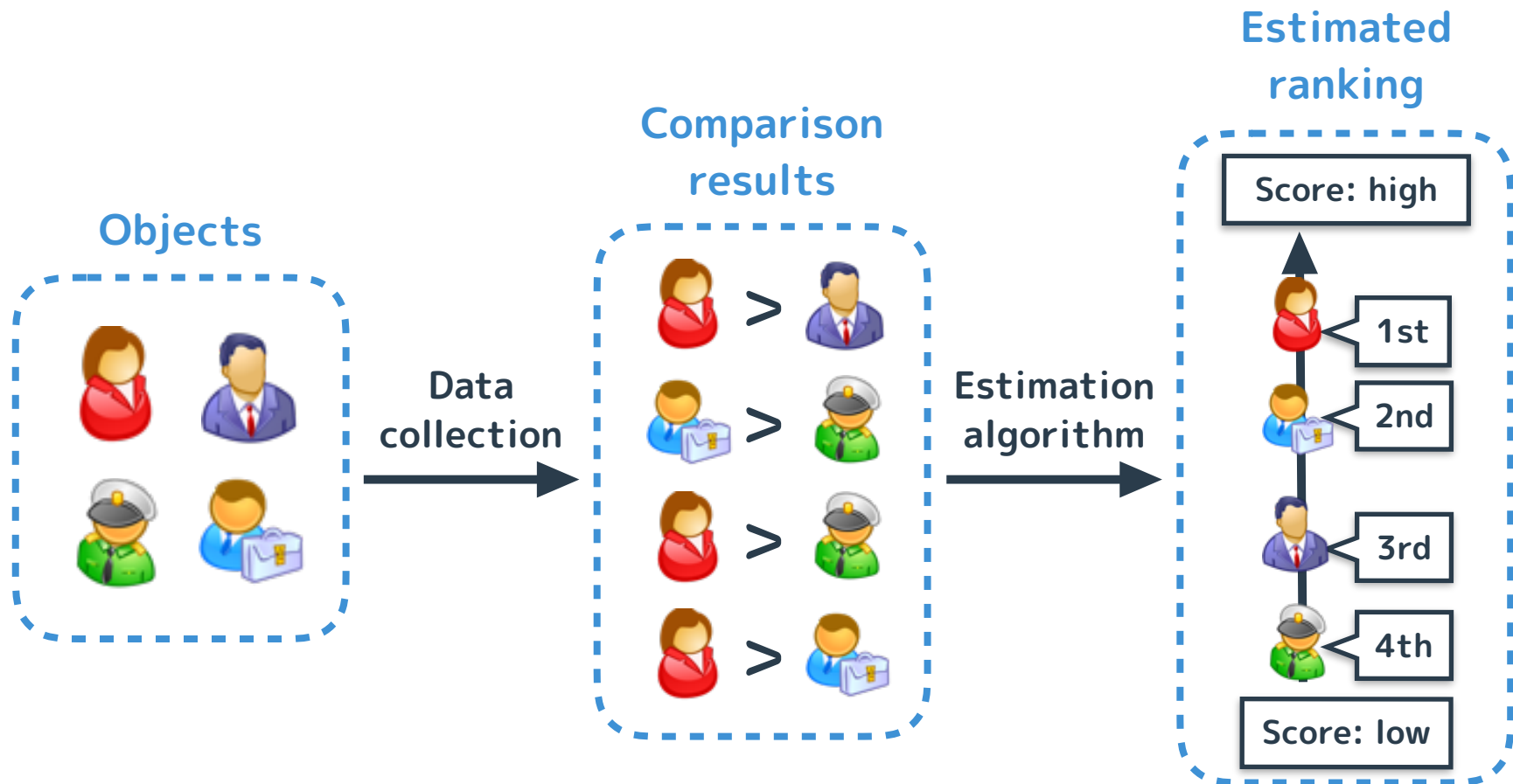
Calculate expectation of changes of matrices

- Expectation of changes in matrices calculated by KL divergence between Bernoulli distributions:



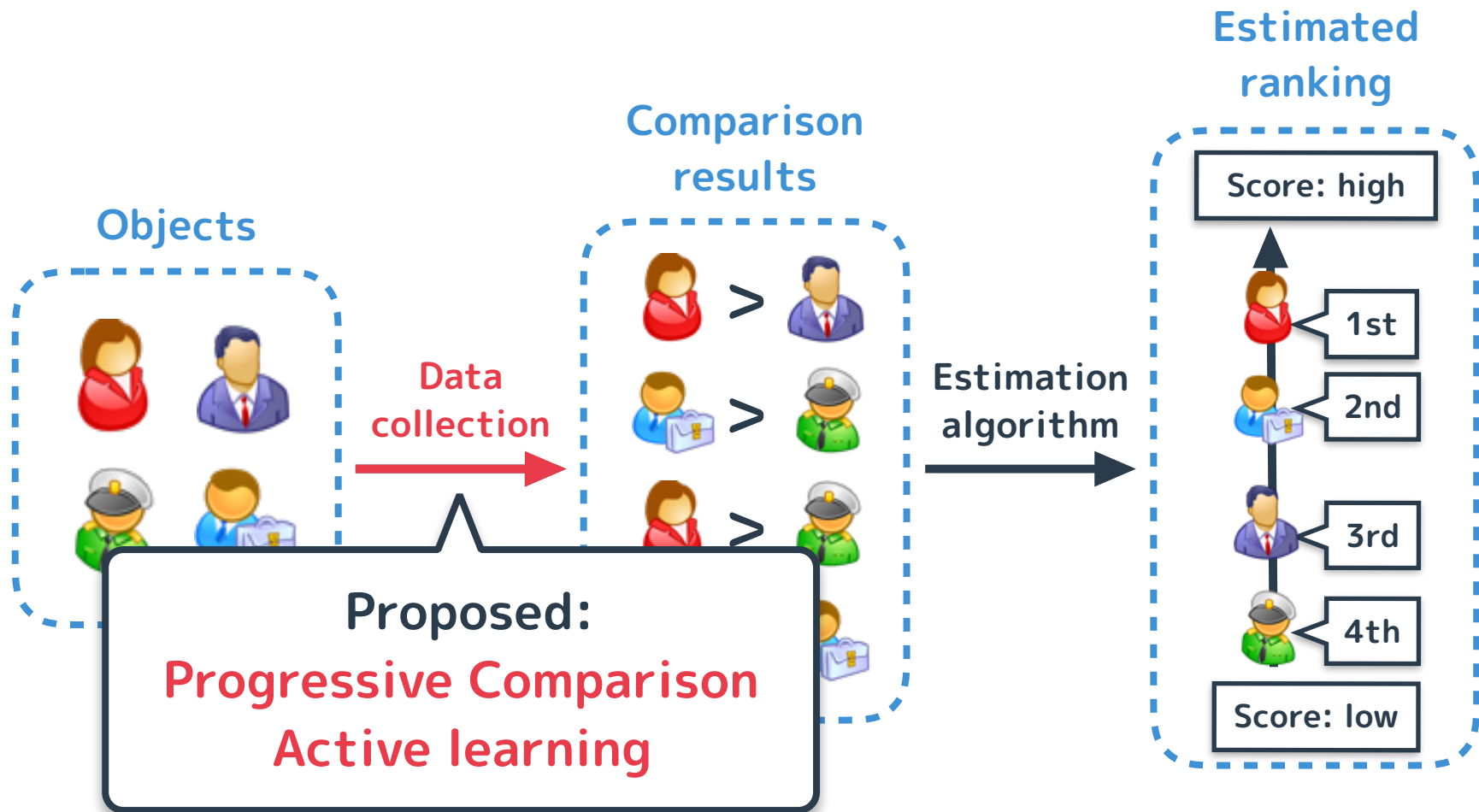
Experiment settings:

Ranking estimation using Glicko Update Equation



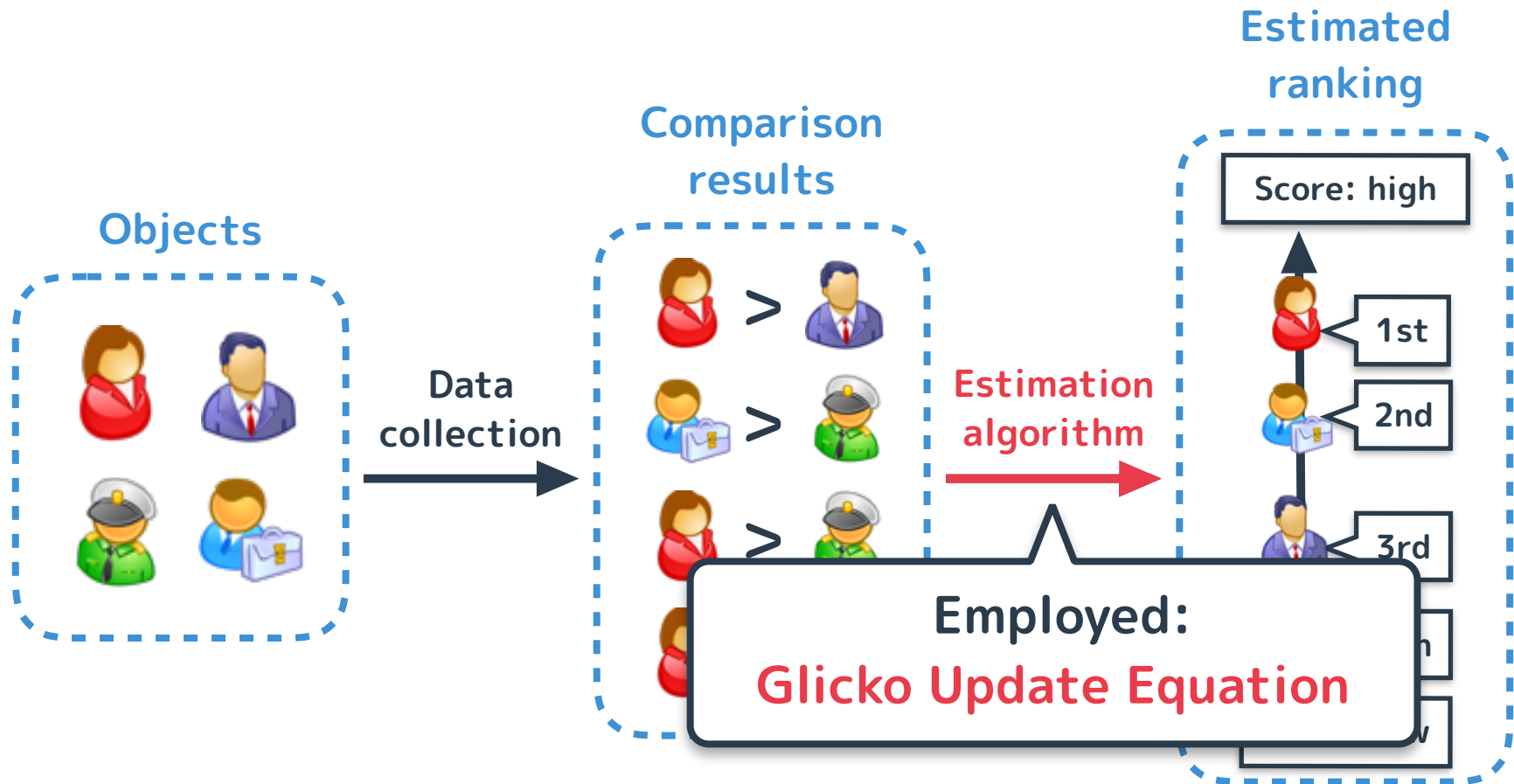
Experiment settings:

Ranking estimation using Glicko Update Equation



Experiment settings:

Ranking estimation using Glicko Update Equation



Experiment settings:

Ranking estimation using Glicko Update Equation

- Glicko Update Equation:
 - Online ranking estimation algorithm of Bradley-Terry model
 - Update scores of object using comparison result

$$\mu' = \mu + \frac{q}{1/\sigma^2 + 1/\delta^2} \sum_{j=1}^m \sum_{k=1}^{n_j} g(\sigma_j^2) \{s_{jk} - E(s|\mu, \mu_j, \sigma_j^2)\},$$

The diagram shows the Glicko Update Equation with three callout boxes. The first box, labeled 'Updated score', points to the symbol μ' . The second box, labeled 'Old score', points to the symbol μ . The third box, labeled 'Comparison result', points to the term s_{jk} in the summation.

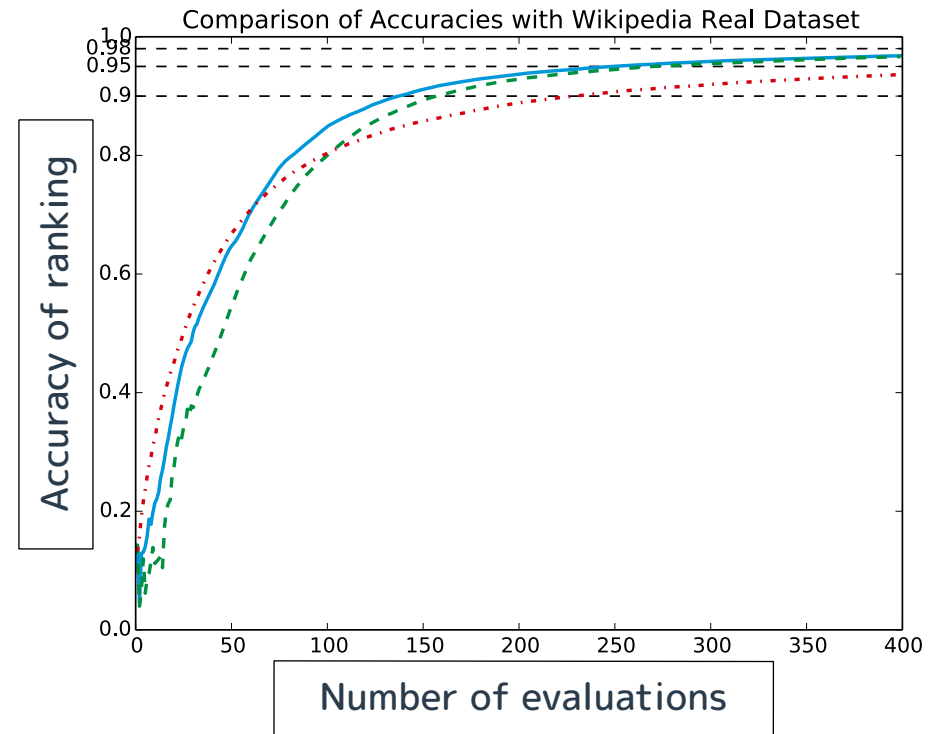
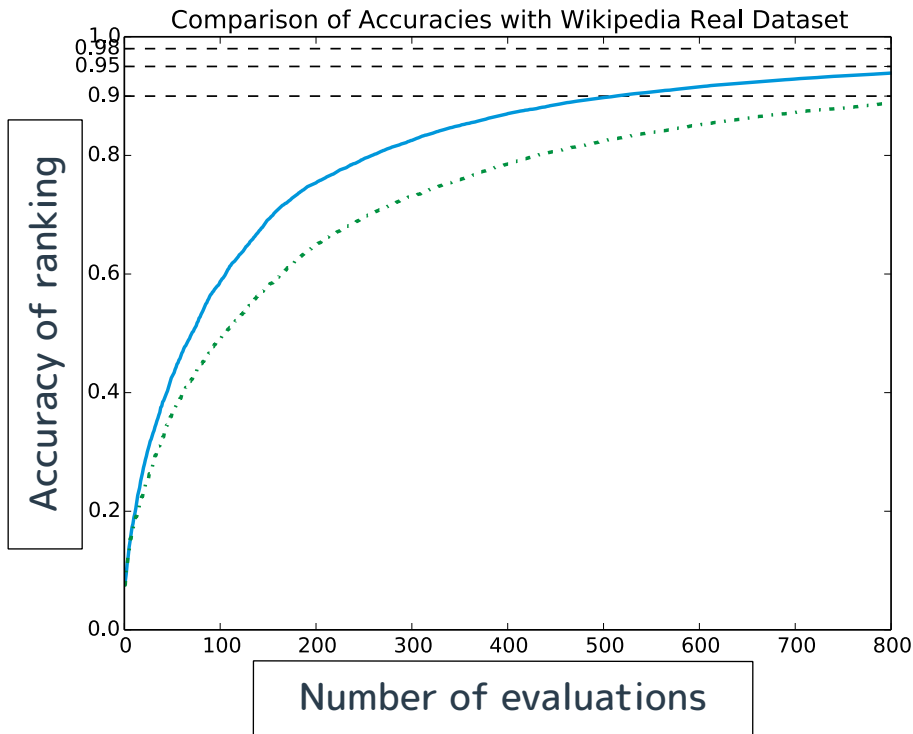
Glickman, Mark E. "Parameter estimation in large dynamic paired comparison experiments." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 48.3 (1999): 377-394.

- Dataset:
 - Synthetic (100 objects)
 - Image comparison (50 objects)
 - Wikipedia article comparison (30 objects)

Experiment results:

Progressive Comparison and active learning methods

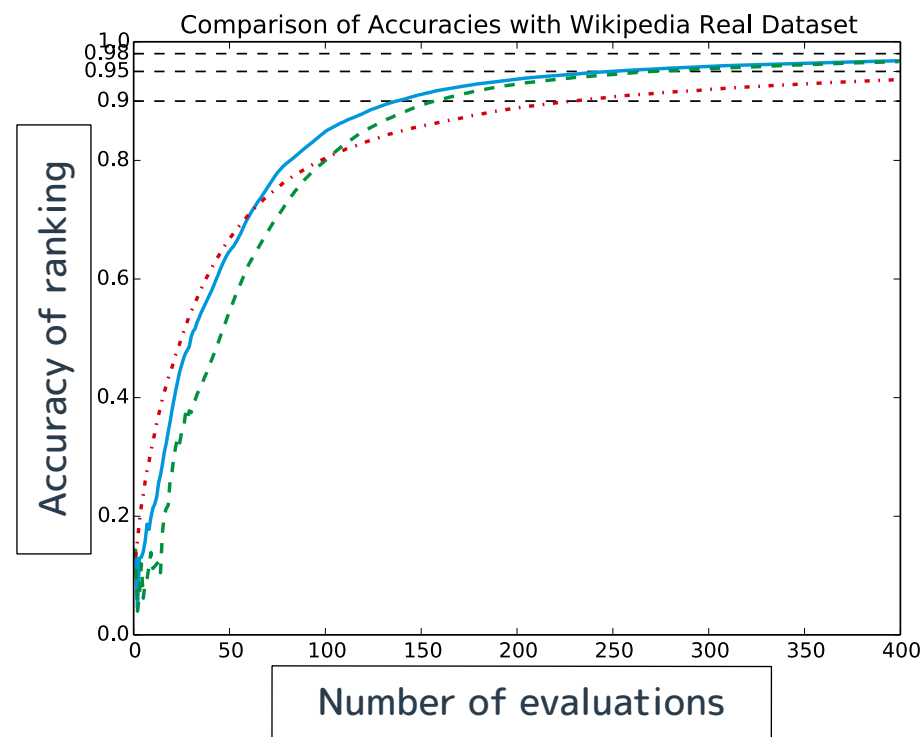
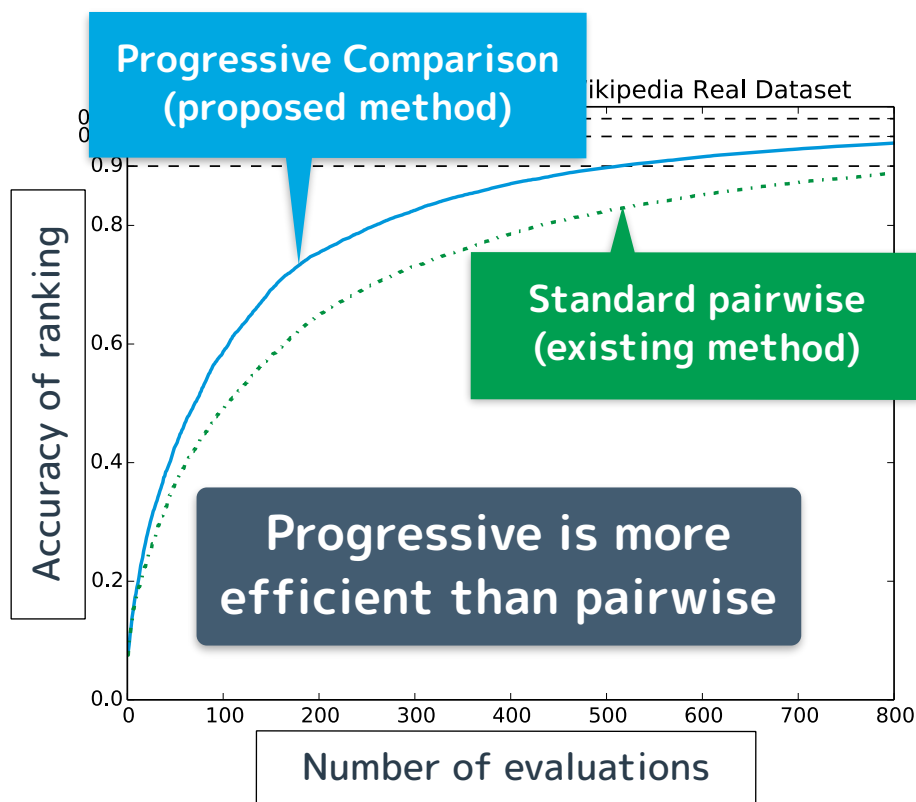
- Experimental results demonstrate the efficiency of Progressive Comparison and its active learning methods



Experiment results:

Progressive Comparison and active learning methods

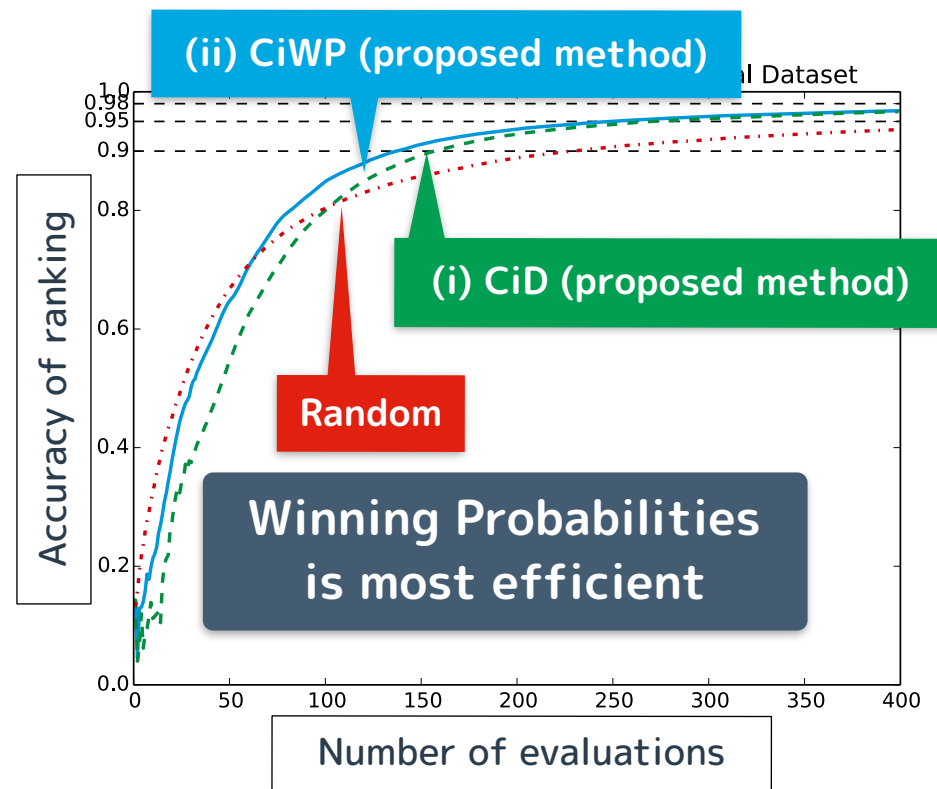
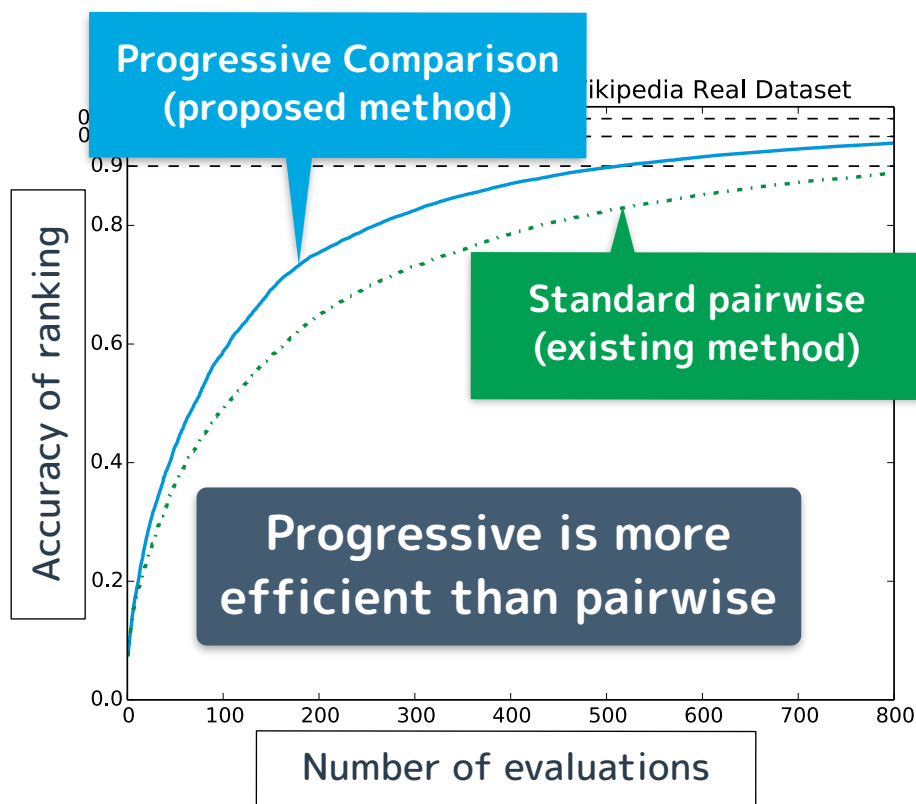
- Experimental results demonstrate the efficiency of Progressive Comparison and its active learning methods



Experiment results:

Progressive Comparison and active learning methods

- Experimental results demonstrate the efficiency of Progressive Comparison and its active learning methods



Conclusions:

Progressive Comparison for Ranking Estimation

- **Ranking estimation** problem addressed
- Proposed:
 - **Progressive comparison**
 - **Active learning** method of Progressive Comparison
 - Change in Distributions
 - Change in Winning Probabilities
- Experimental results show:
 - Superiority of Progressive Comparison to standard pairwise
 - Efficiency of active learning methods for Progressive Comparison (especially (ii) CiWP)