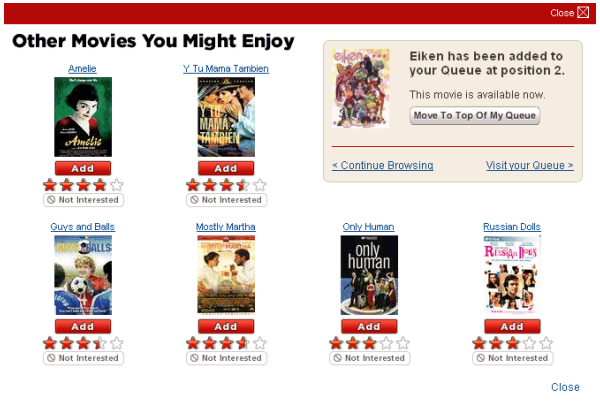


Online Learning to Rank with Features

Authors: Shuai Li, Tor Lattimore, Csaba Szepesvári

The Chinese University of Hong Kong
DeepMind
University of Alberta

Learning to Rank



Amazon, YouTube, Facebook, Netflix, Taobao

Online Learning to Rank

- There are $\hat{I}$ items and $\hat{E}$ $\hat{I}$ positions
- At each time $^{?} = 1; 2; : : :$,
  - Choose an ordered list $_{?} = (\mathcal{E}_1^{?}; : : : ; \mathcal{E}_{\hat{E}}^{?})$
  - Show the user the list
  - Receive click feedback $G_{?1}; : : : ; G_{?\hat{E}}$ 2 $f0; 1g$, per position
- Objective: Maximize the expected number of clicks

$$\mathrm{E}\left[\sum_{?=1}^{\hat{}} \sum_{\zeta=1}^{\hat{E}} G_{?\zeta}\right]$$

Click Models

- Click models describe how users interact with item lists
- Cascade Model (CM)
  - Assumes the user checks the list from position 1 to position $\hat{E}$, clicks at the first satisfying item and stops
- Dependent Click Model (DCM)
  - Further assumes there is a satisfaction probability after click
- Position-Based Model (PBM)
  - Assumes the user click probability on an item a of position k can be factored into item attractiveness and position bias
- Generic model
  - Make as few assumptions as possible about the click model

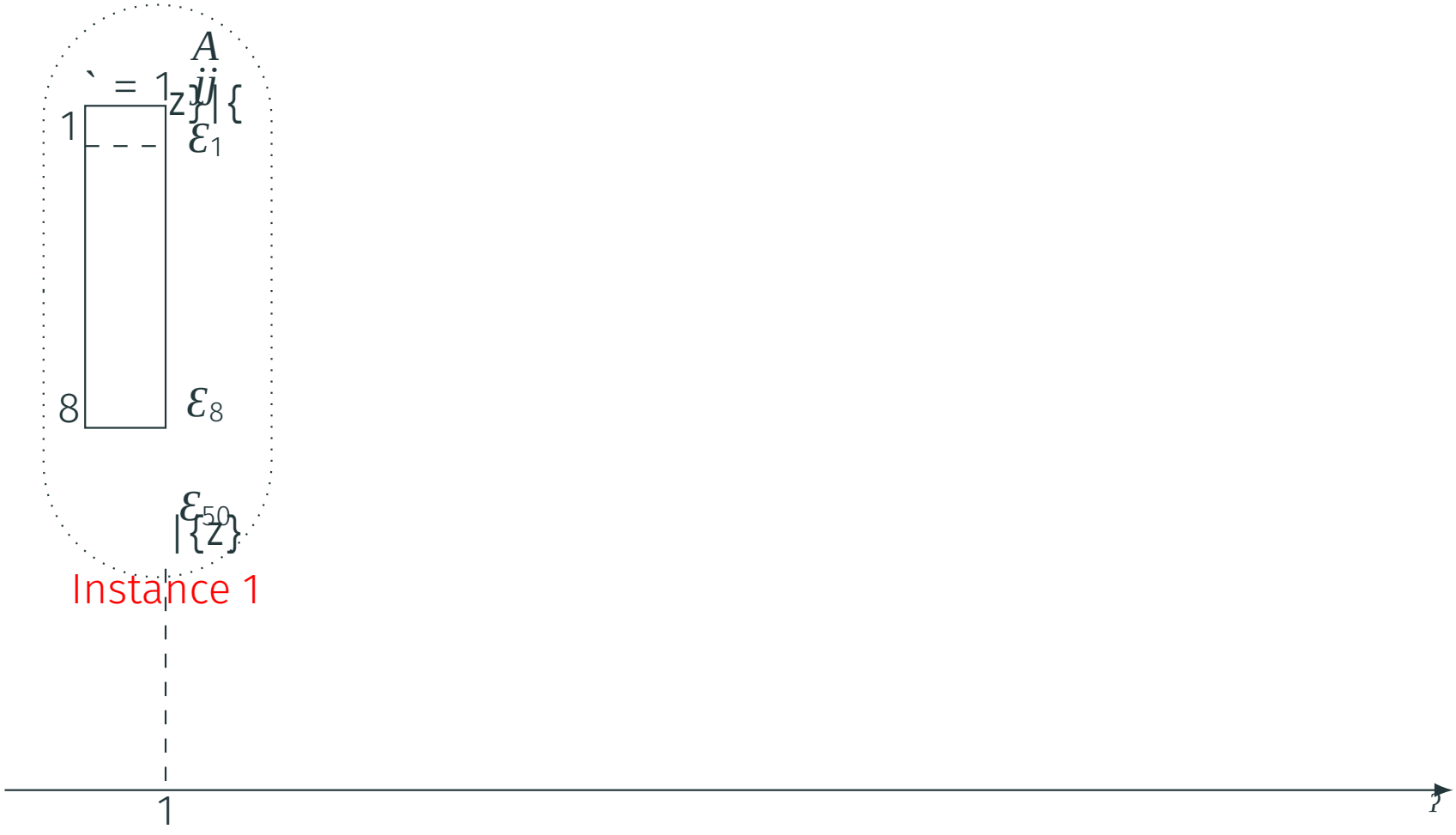
✗
✓
✗
✓
✗

RecurRank

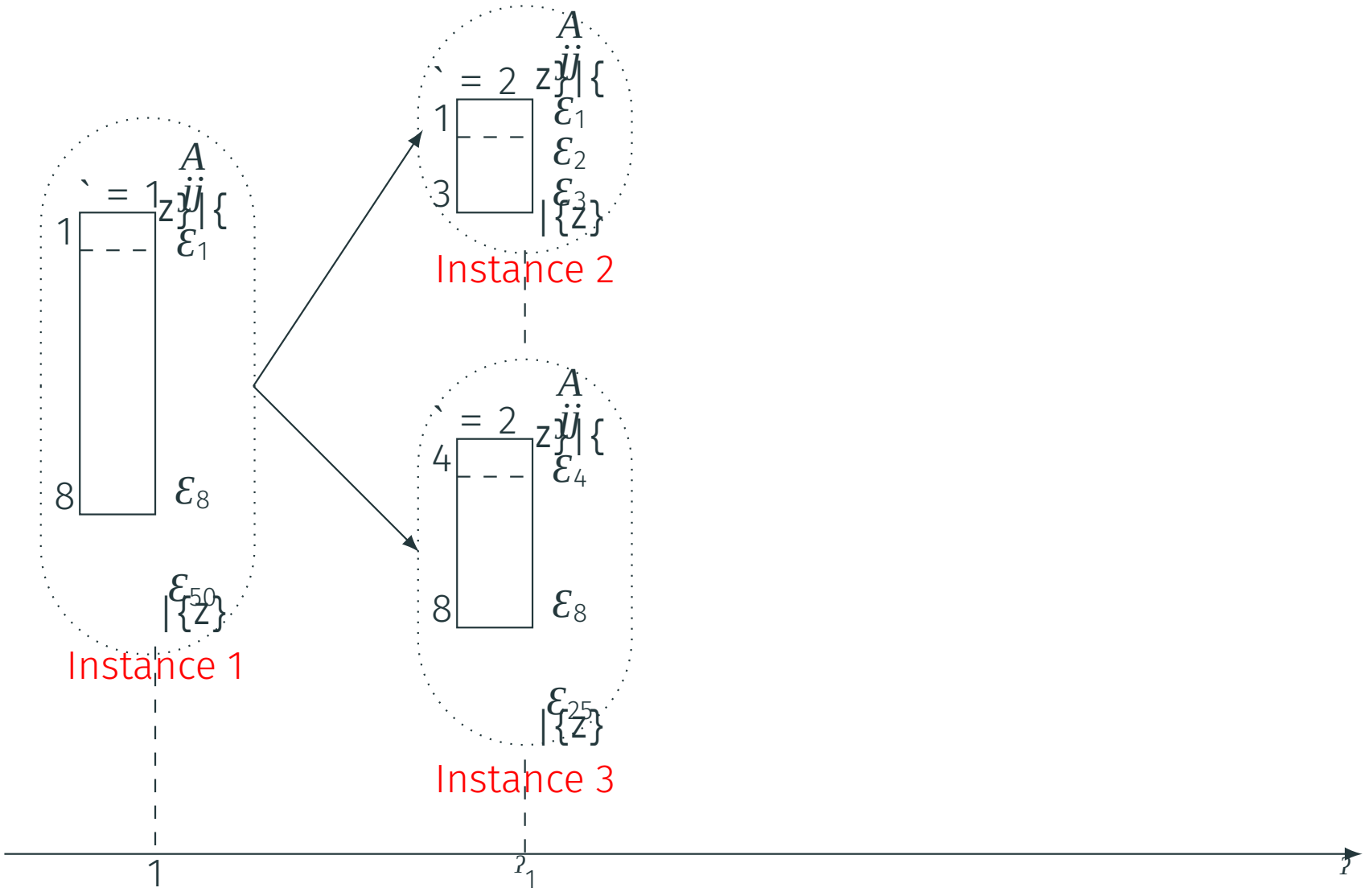
- Each item $\mathcal{E}$ is represented by a feature vector $\llcorner_{\mathcal{E}}\ 2\ \mathbf{R}^{Lj}$
- The attractiveness of item $\mathcal{E}$ is $(\mathcal{E}) = {}^{\top}\llcorner_{\mathcal{E}}$
- Click probability factors: $\mathbf{P}_{?}(G_{?\zeta} = 1) = (\mathcal{E}^{?}_{\zeta})\ (\ _{?};\zeta)$ where is the examination probability, which satisfies reasonable assumptions

- RecurRank (Recursive Ranking)
- For each phase `
  - Use first position for exploration
  - Use remaining positions for exploitation, rank best items first
- Split items and positions when the phase ends
- Recursively call the algorithm with increased phase

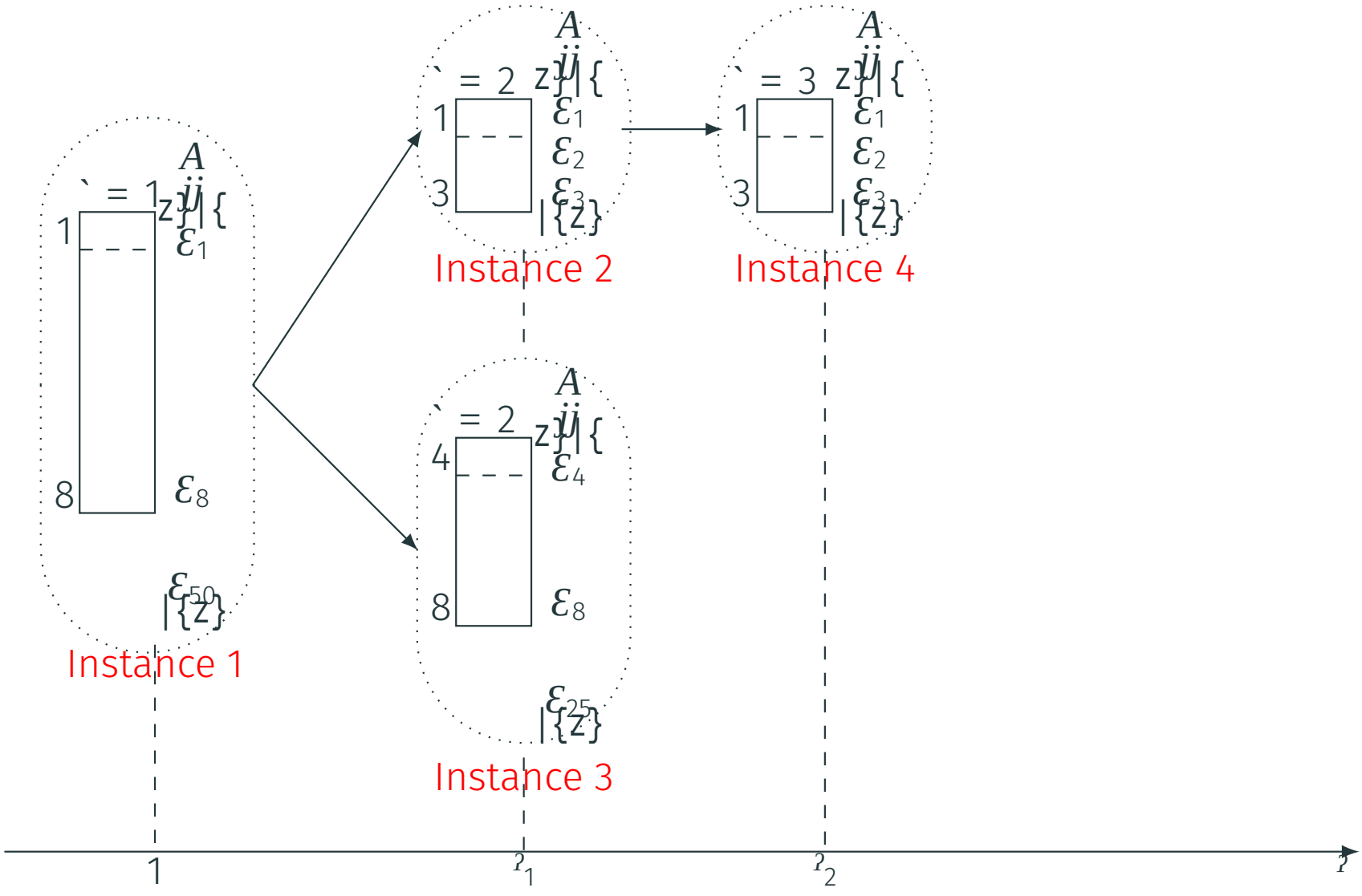
Example



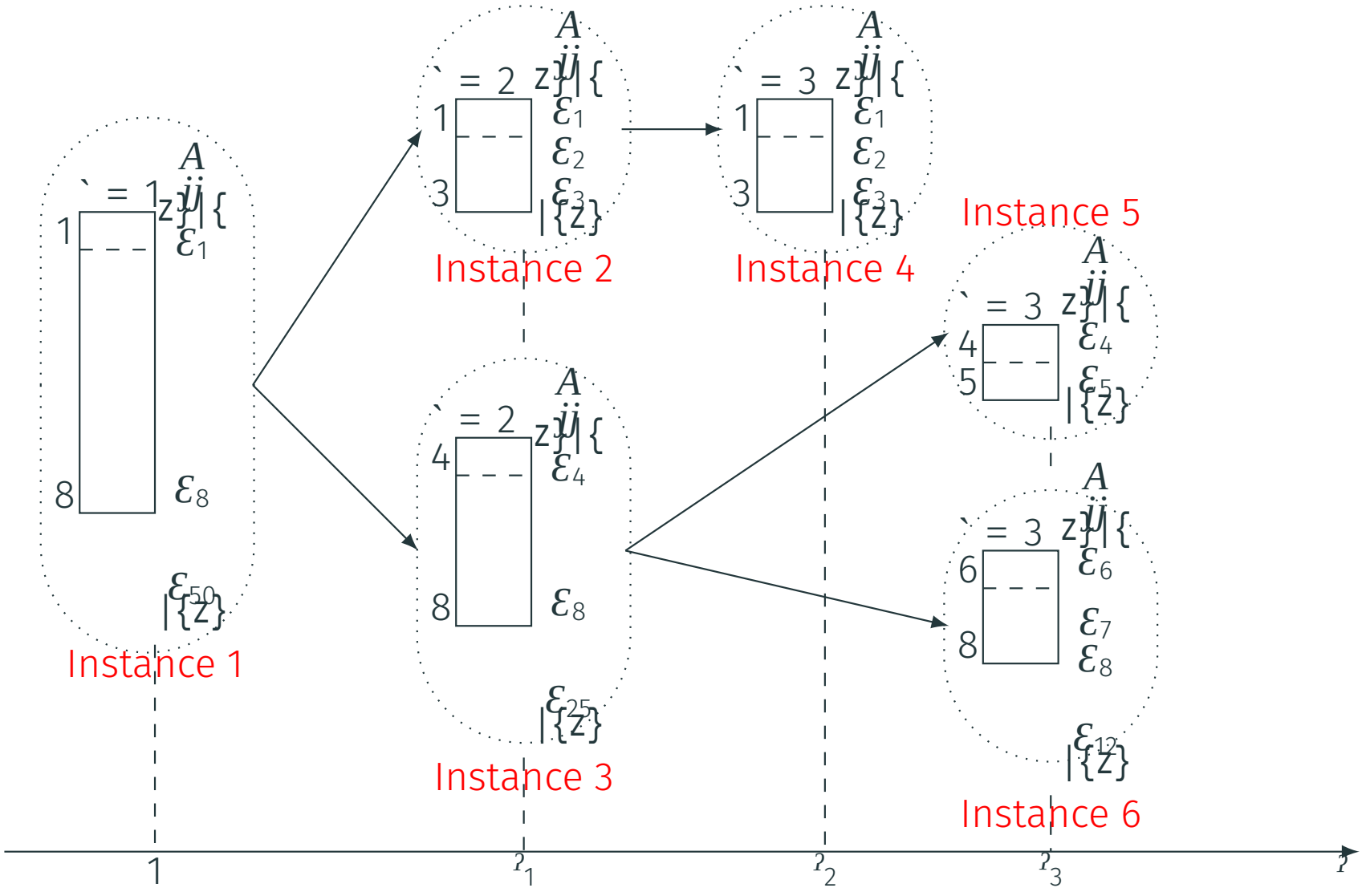
Example



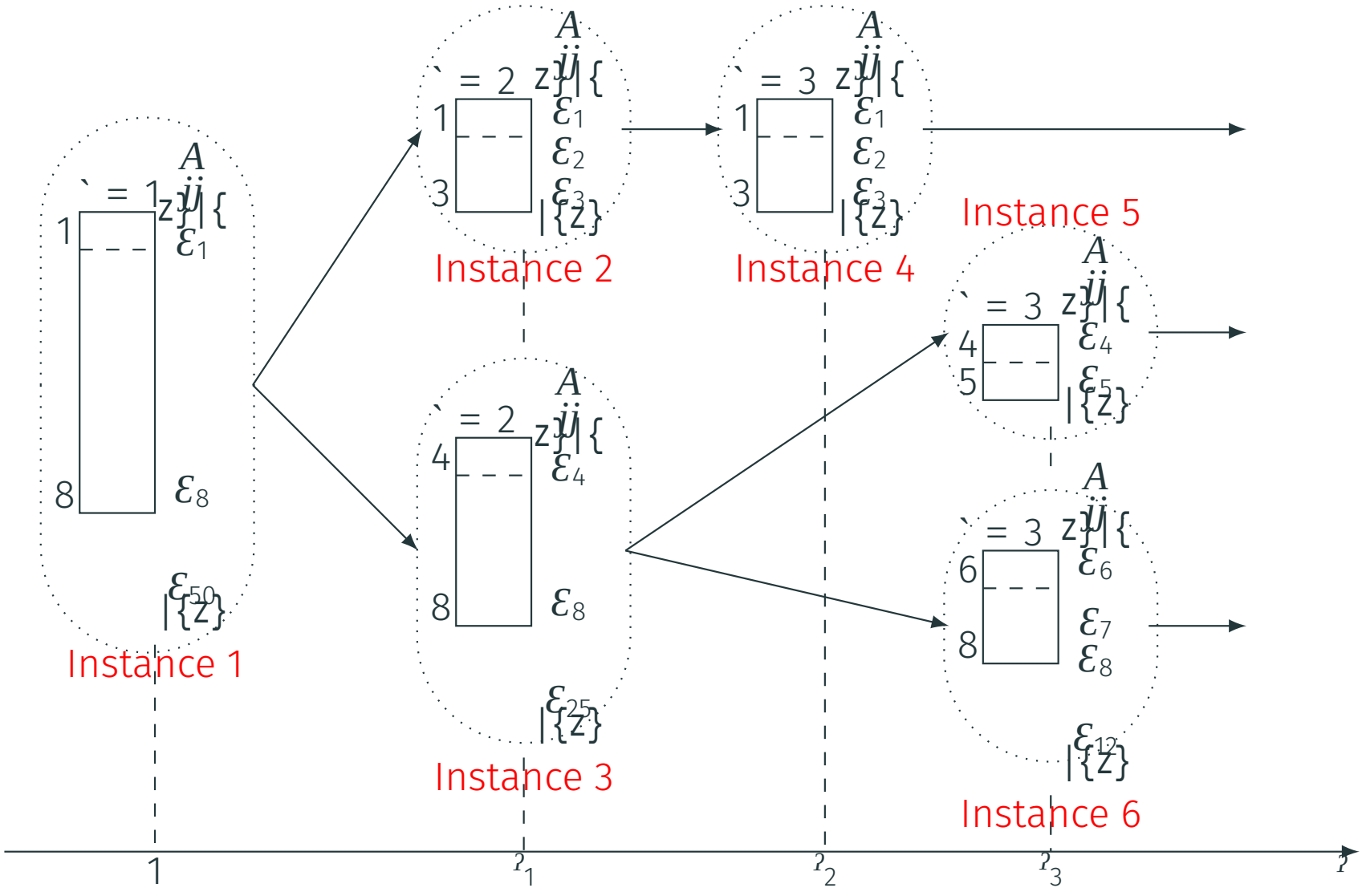
Example



Example


Instance 1
Instance 2
Instance 3
Instance 4
Instance 5
Instance 6
1
r_1
r_2
r_3
t

Example


Instance 1
Instance 2
Instance 3
Instance 4
Instance 5
Instance 6

Results

- Regret bound

$$\hat{g}(\imath) = \grave{\imath}\,(\hat{E}\,\sqrt{Lj\imath\,\log(\hat{I}\,\imath)})$$

- Improves over existing bound $\grave{\imath}\,\sqrt{\hat{E}^3\hat{I}\,\imath\,\log(\imath)}$

Results

- Regret bound

$$\hat{g}(\imath) = \grave{\imath}\,(\hat{E}\sqrt{Lj\imath\log(\hat{I}\imath)})$$

- Improves over existing bound $\grave{\imath}\sqrt{\hat{E}^3\hat{I}\imath\log(\imath)}$

—RecurRank —CascadeLinUCB —TopRank

Thank you!

Sumeet Katariya, Branislav Kveton, Csaba Szepesvari, and Zheng Wen.
Dcm bandits: Learning to rank with multiple clicks.
In *[illegible]*, pages 1215–1224, 2016.

Branislav Kveton, Csaba Szepesvari, Zheng Wen, and Azin Ashkan.
Cascading bandits: Learning to rank in the cascade model.
In *[illegible]*, pages 767–776, 2015.

Paul Lagrée, Claire Vernade, and Olivier Cappe.
Multiple-play bandits in the position-based model.
In *[illegible]*, pages 1597–1605, 2016.

References ii

Tor Lattimore, Branislav Kveton, Shuai Li, and Csaba Szepesvari.
Toprank: A practical algorithm for online stochastic ranking.
In *ı ṣÙ GƷɕâÙȡÙɕʂÙ Ʒɕ ÞÙ·ȡƐǼ ³ɕâƷȡ ɔ Ɛ²ɀƷɕ ĖȡƷʂÙʳ ʳɀɕÊ Ħ_ʳ²Ù ɔ ʳ*, 2018.

Shuai Li, Tor Lattimore, and Csaba Szepesvári.
Online learning to rank with features.
Ɛȡtɀ¸ ʙȡÙʙȡɀɕ² Ɛȡtɀ¸╣ ⁶┬┼ ╣┃┼ ⁴⁵, 2018.

Shuai Li, Baoxiang Wang, Shengyu Zhang, and Wei Chen.
Contextual combinatorial cascading bandits.
In *³ɕ²ÙȡɕƐ²ɀƷɕƐǼ GƷɕâÙȡÙɕʂÙ Ʒɕ ÚƐʂṣɀÙ ÎÙƐȡɕɀÊ*, pages 1245–1253, 2016.

Shuai Li and Shengyu Zhang.
Online clustering of contextual cascading bandits.
In *ı ṣÙ ³ GƷɕâÙȡÙɕʂÙ Ʒɕ ȡ²ɀ ʂɀƐǼ ³ɕ²ÙǼǼɀÊÙɕʂÙ*, 2018.

Weiwen Liu, Shuai Li, and Shengyu Zhang.
Contextual dependent click bandit algorithm for web recommendation.
In *ʒɞʔỦȡɞƐʔɿʒɞƐȽ Gʒɘ ʙ ˑʔɿɞÊ ƐɞǇ Gʒɘ ʏɿɞƐʔʒȡɿʂ ʳ GʒɞâỦȡỦɞʂỦ*, pages 39–50. Springer, 2018.

Masrour Zoghi, Tomas Tunys, Mohammad Ghavamzadeh, Branislav Kveton, Csaba Szepesvari, and Zheng Wen.
Online learning to rank in stochastic click models.
In *ʒɞʔỦȡɞƐʔɿʒɞƐȽ GʒɞâỦȡỦɞʂỦ ʒɞ ÚƐʂșɿɞỦ ÎỦƐȡɞɿɞÊ*, pages 4199–4208, 2017.

Shi Zong, Hao Ni, Kenny Sung, Nan Rosemary Ke, Zheng Wen, and Branislav Kveton.
Cascading bandits for large-scale recommendation problems.
In *[illegible]*, pages 835–844. AUAI Press, 2016.