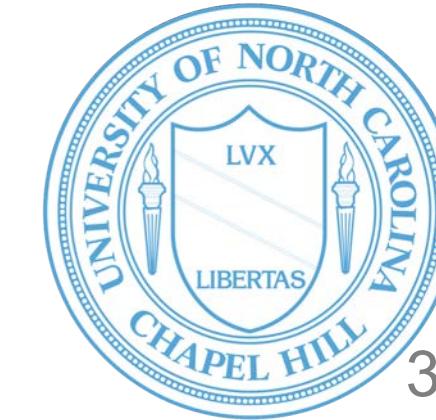


Visual Progression Analysis of Event Sequence Data

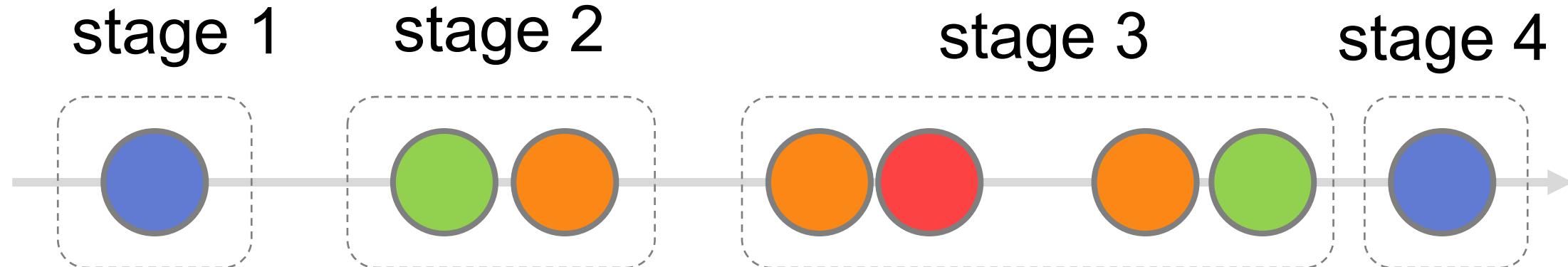
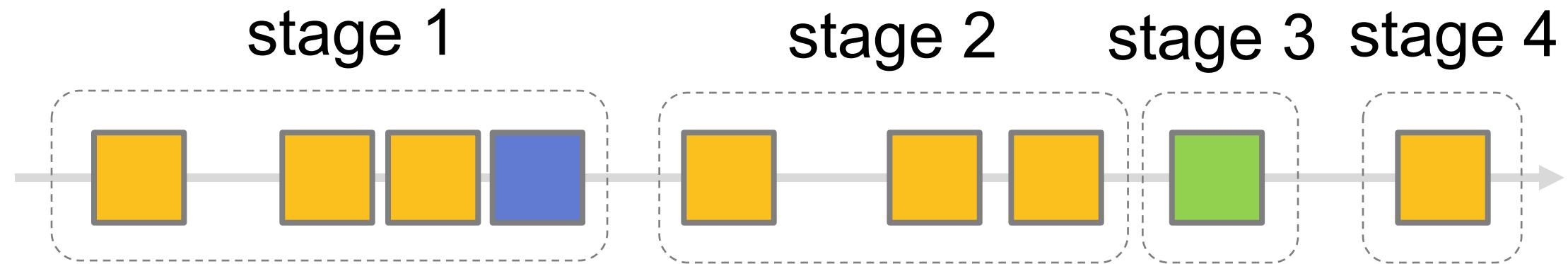
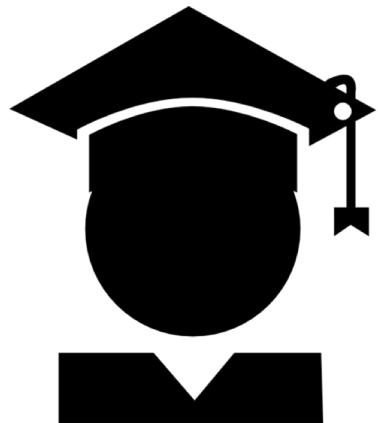
Shunan Guo¹, Zhuochen Jin², David Gotz³, Fan Du⁴, Hongyuan Zha¹, Nan Cao²✉



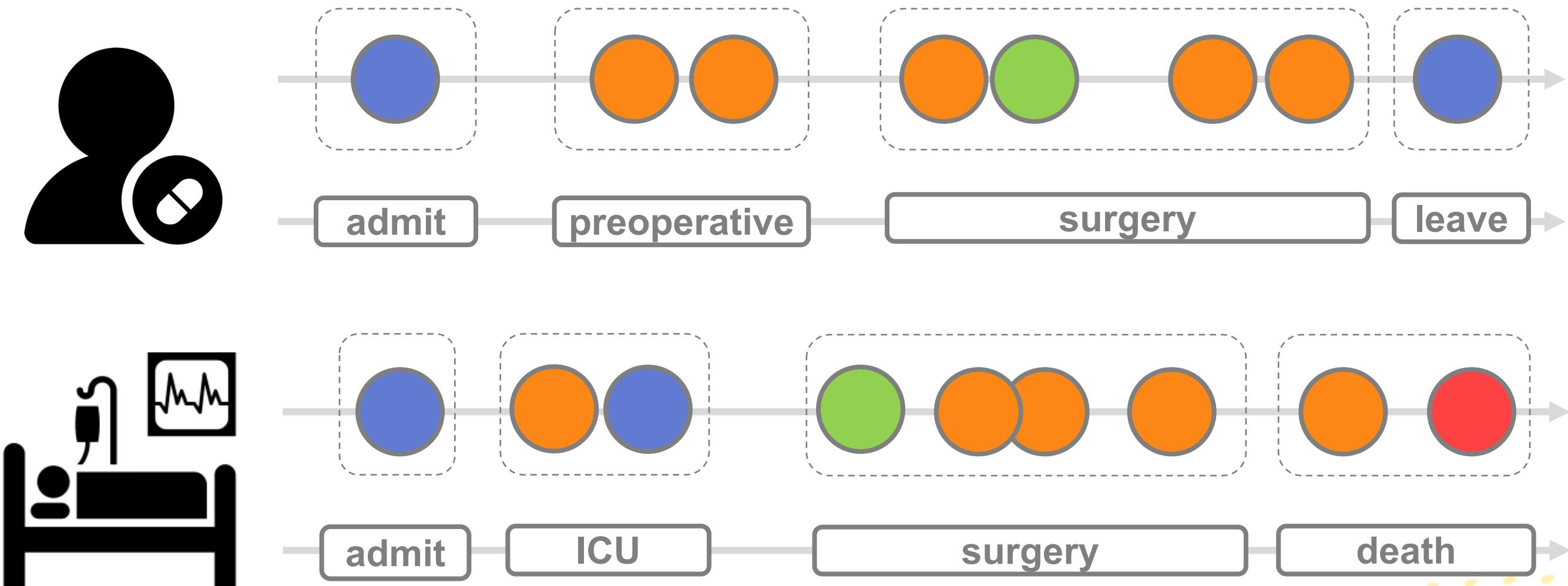
Visual Progression Analysis of Event Sequence Data

Shunan Guo¹, Zhuochen Jin², David Gotz³, Fan Du⁴, Hongyuan Zha¹, Nan Cao²✉

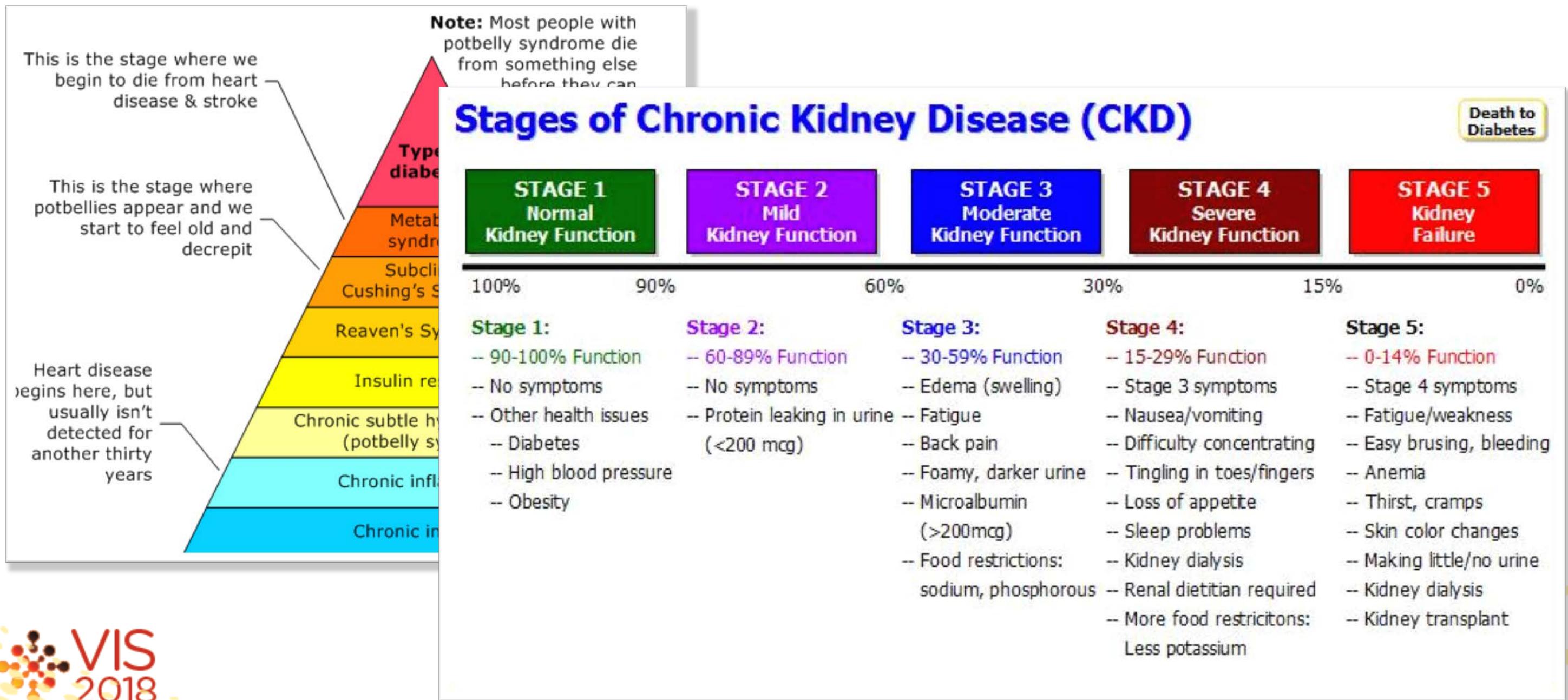
Progression Analysis on Event Sequences



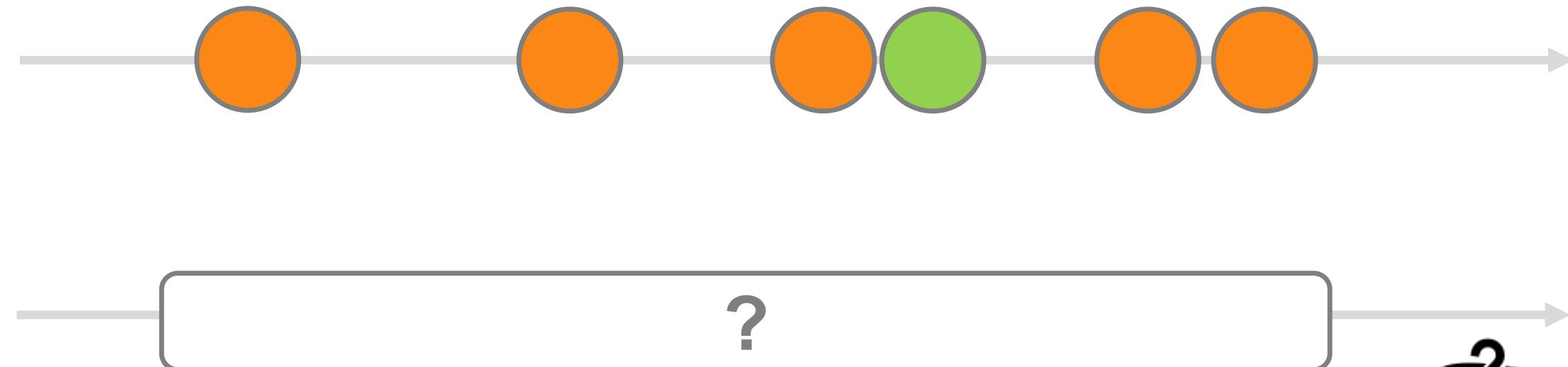
Progression of Medical Procedures



Knowing Progression Patterns is Useful



Challenges in Progression Analysis



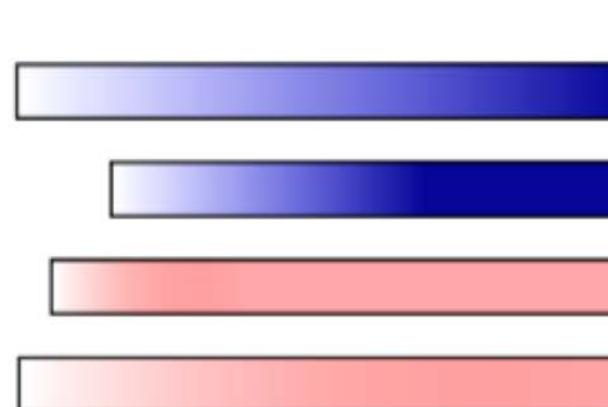
Different length / progression rate / stages



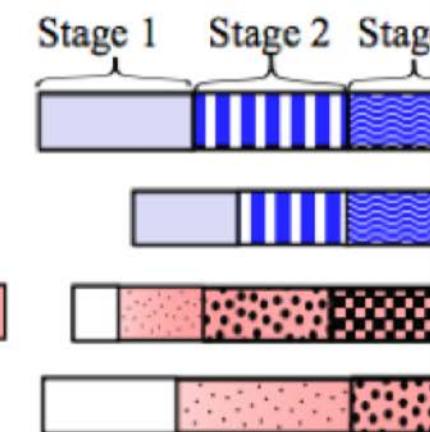
Challenges in Progression Analysis

Finding Progression Stages in Time-evolving Event Sequences

Jaewon Yang^{†*} Julian McAuley[†] Jure Leskovec[†] Paean Ling[‡]
† Computer Science, Stanford University, {jayang, jmcauley, jure}@cs.stanford.edu
‡ Biomedical Informatics, Stanford University, {plependu, nihal}@stanford.edu



(a) Input sequences



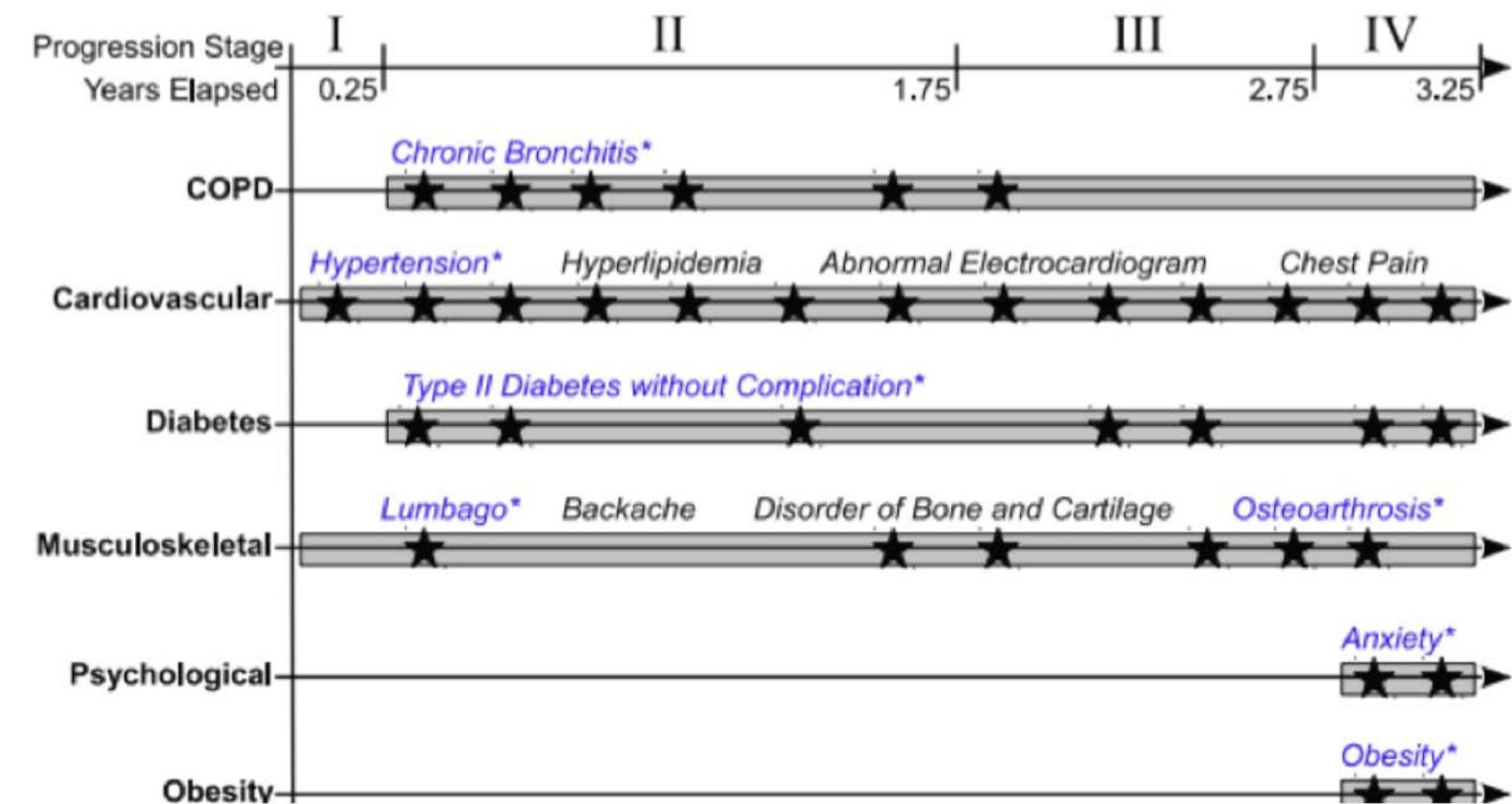
(b) Stages

Unsupervised Learning of Disease Progression Models

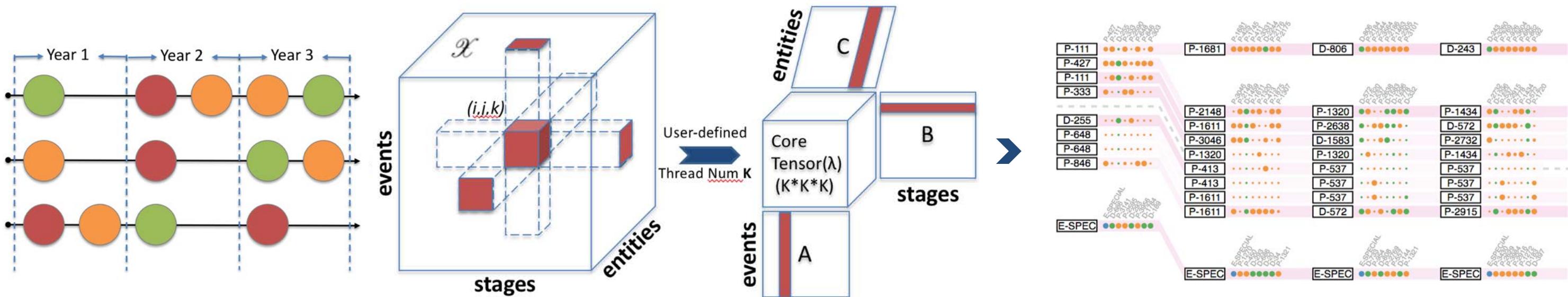
Xiang Wang
IBM Research
Yorktown Heights, NY
wangxi@us.ibm.com

David Sontag
New York University
New York, NY
dsontag@cs.nyu.edu

Fei Wang
IBM Research
Yorktown Heights, NY
fwang@us.ibm.com



Previous Study - EventThread



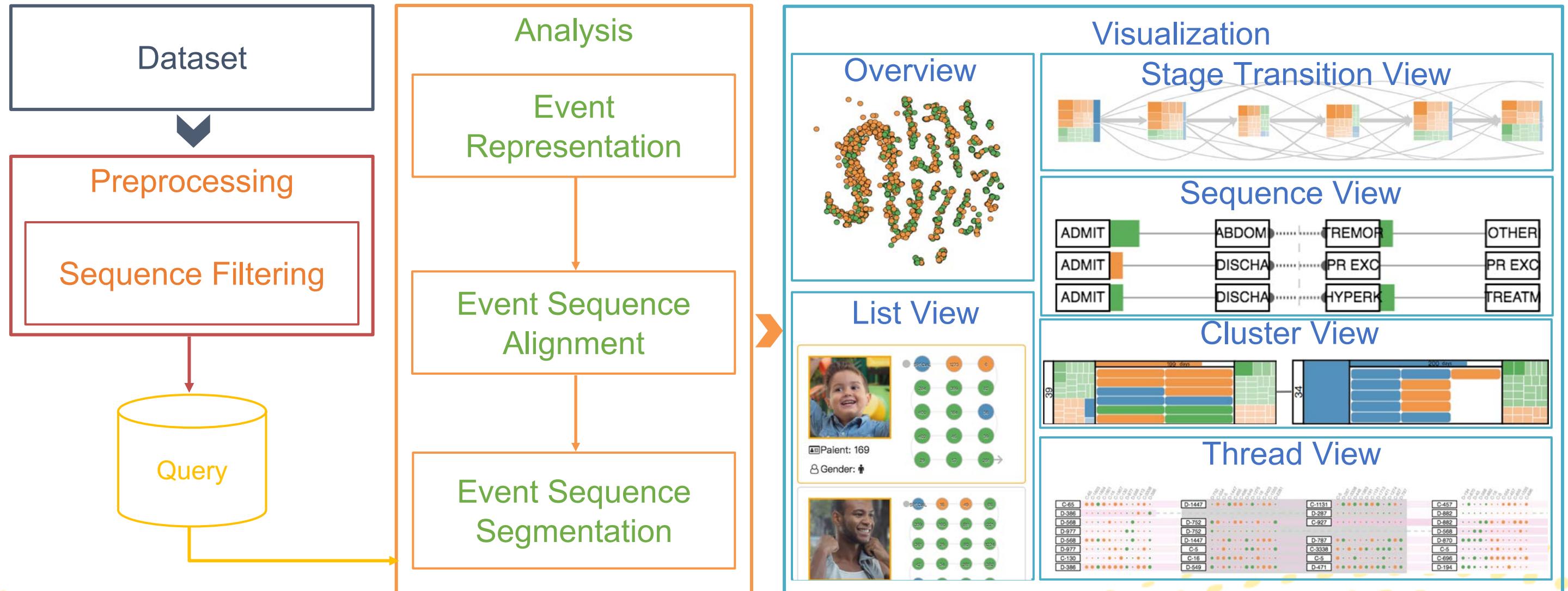
Guo, S., Xu, K., Zhao, R., Gotz, D., Zha, H., & Cao, N. (VAST'2018).

EventThread: Visual Summarization and Stage Analysis of Event Sequence Data.

We introduce EventThread2

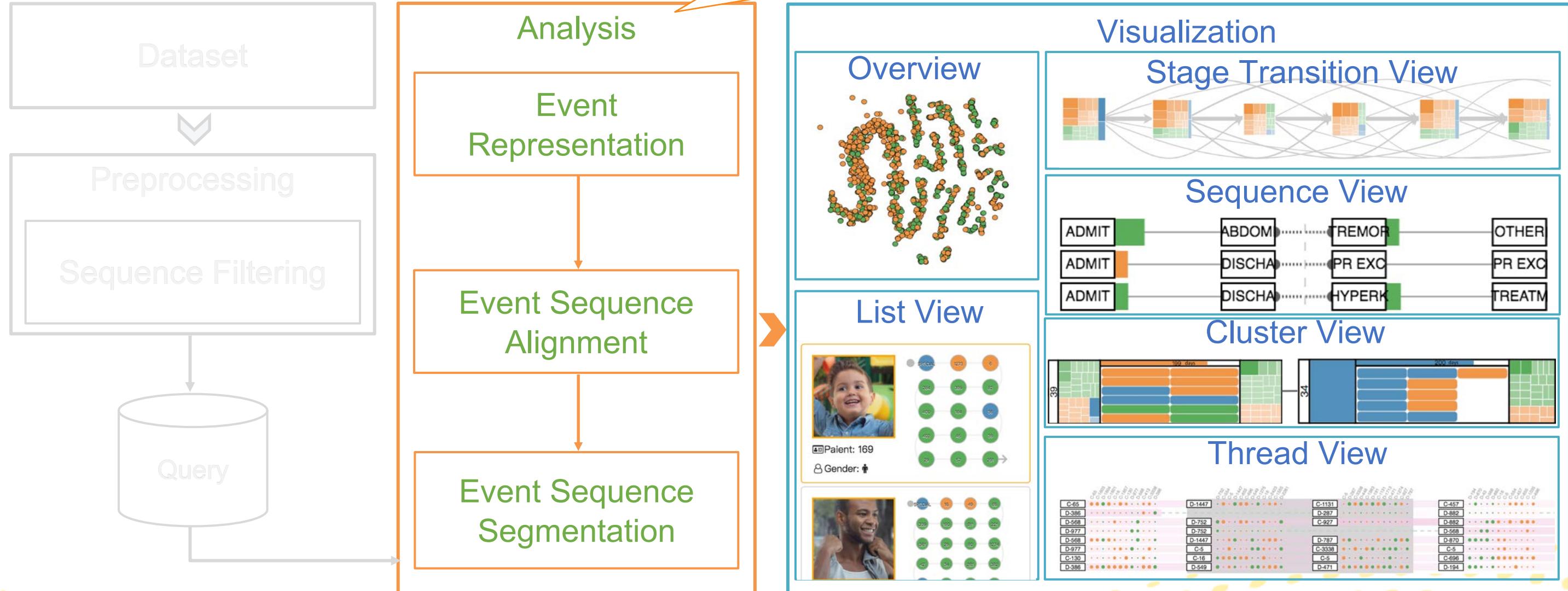


System Overview

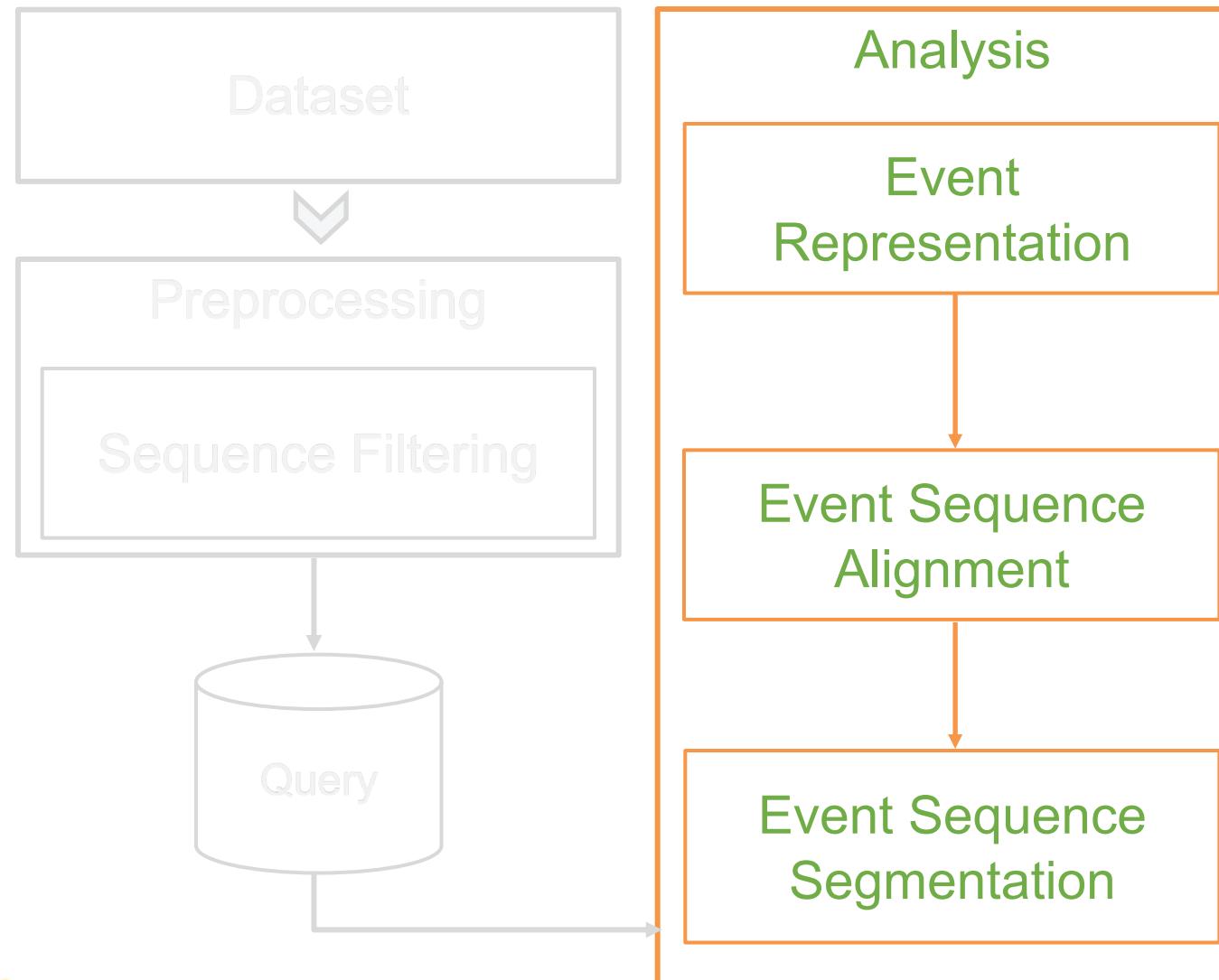


System Overview

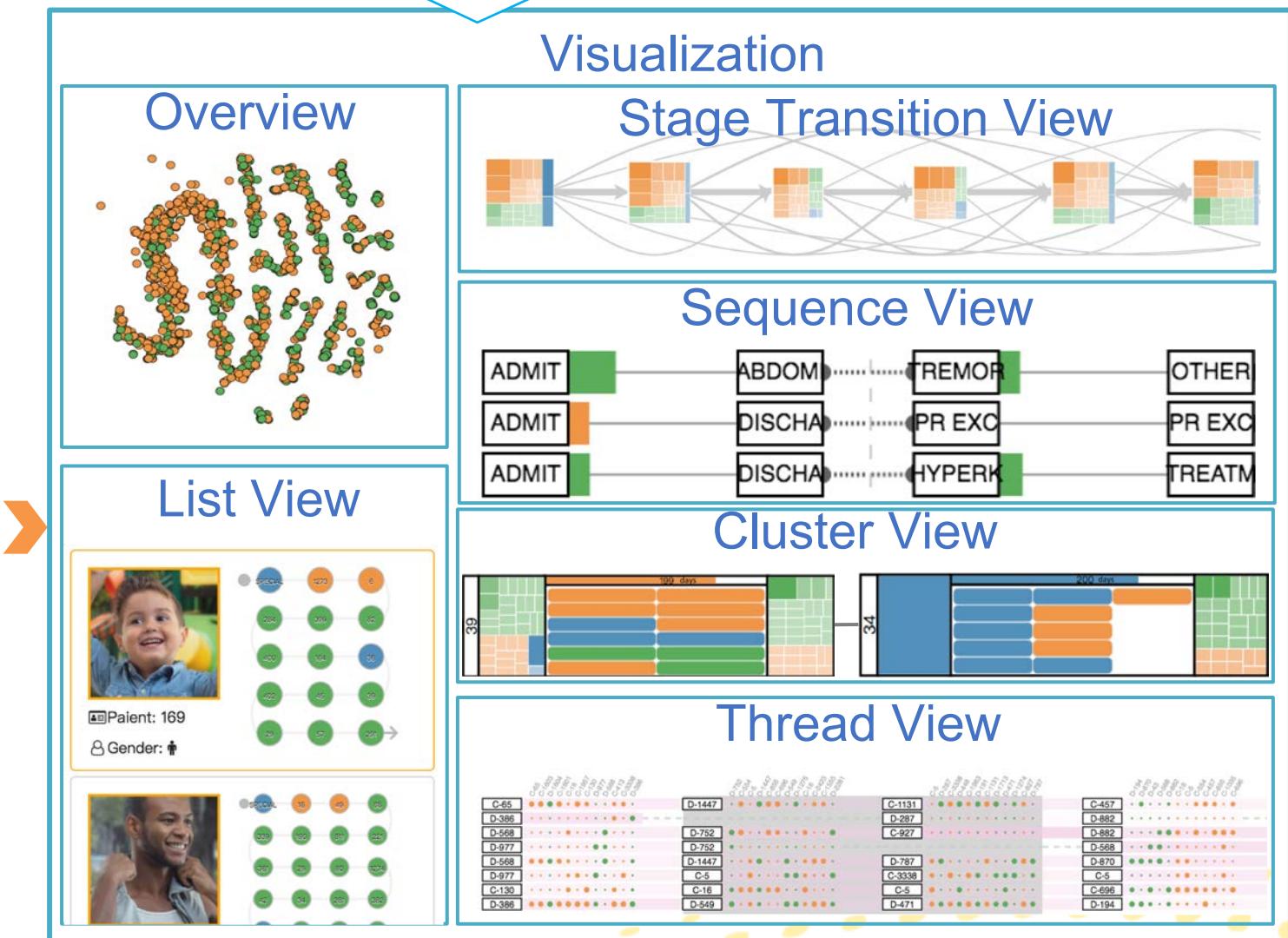
How to identify progression stages of a sequence set?



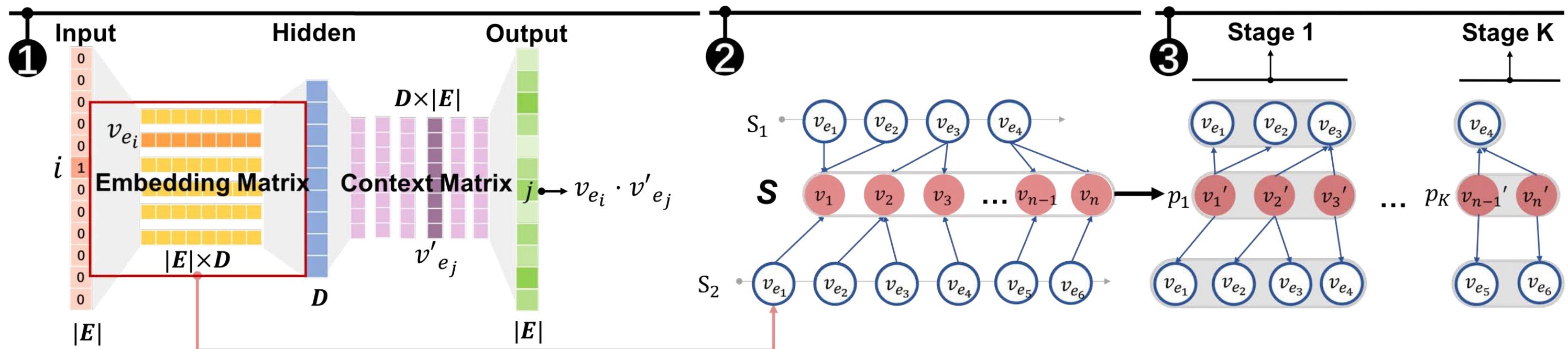
System Overview



How to visualize the result?

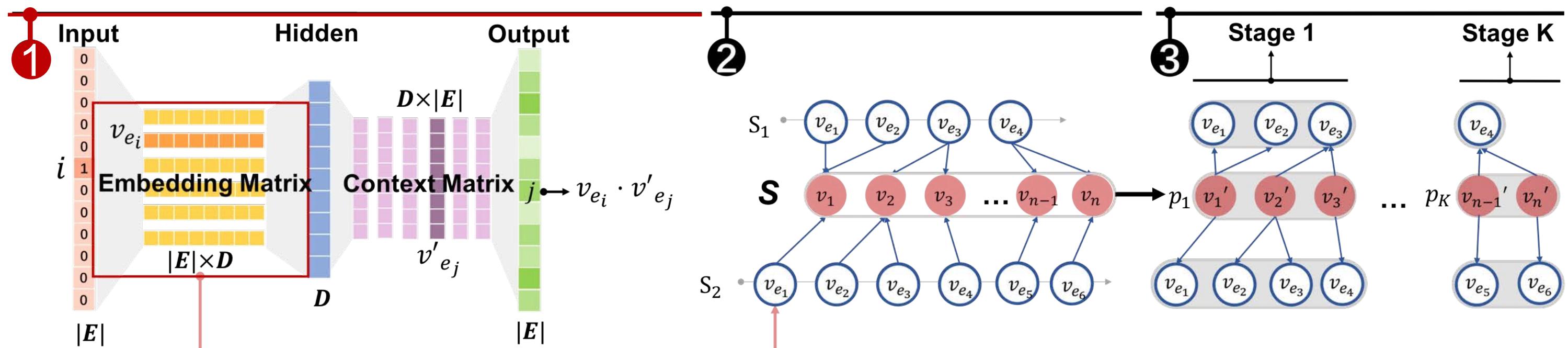


Progression Analysis

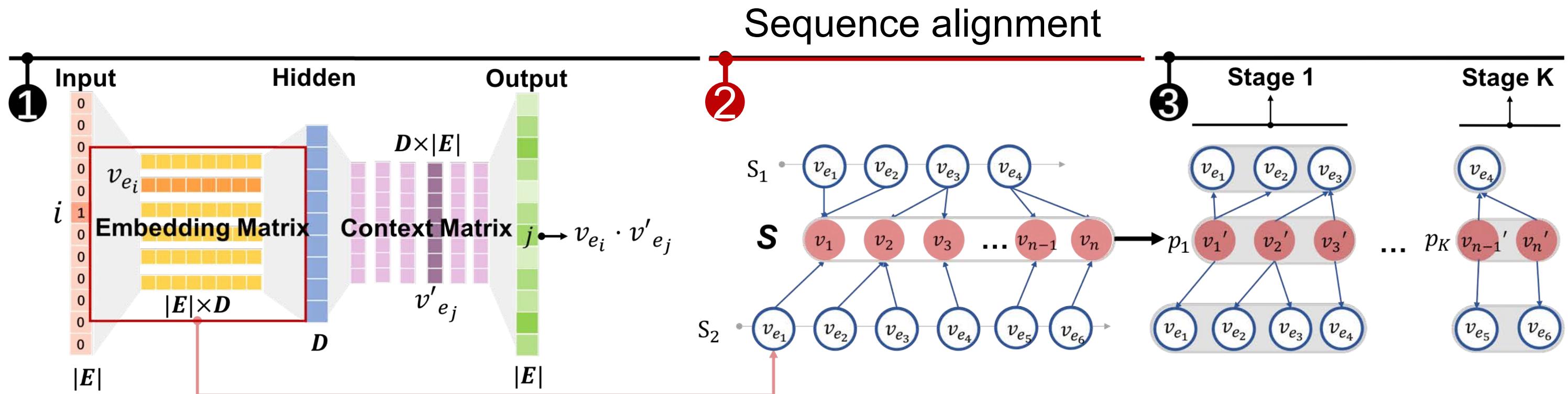


Progression Analysis

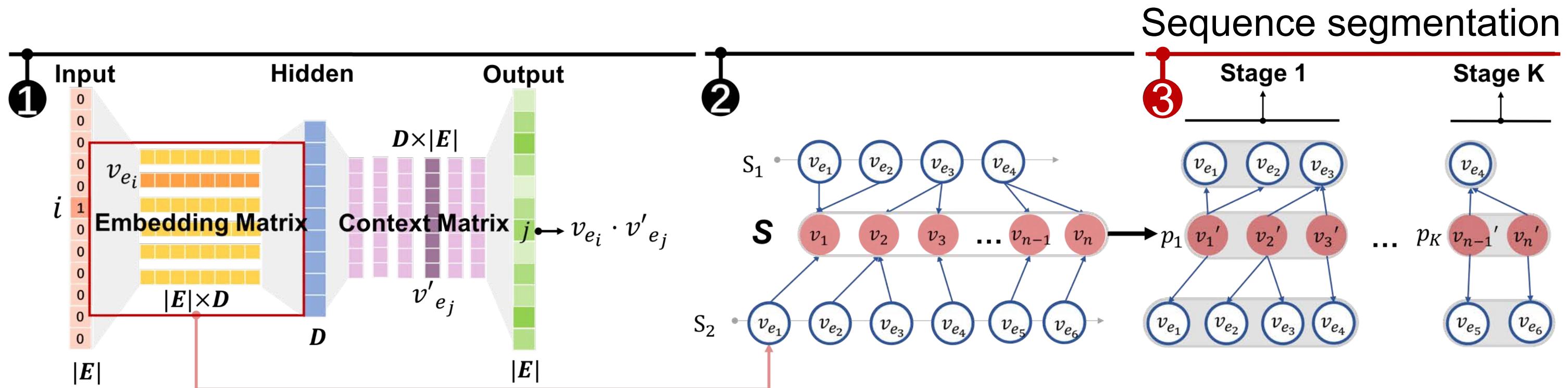
Estimation of event representation



Progression Analysis



Progression Analysis

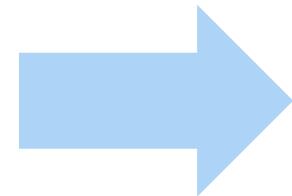


Estimate Event Representations

Which session will you go next



Sentence



Phrases

Estimate Event Representations

Which session

will you go

next

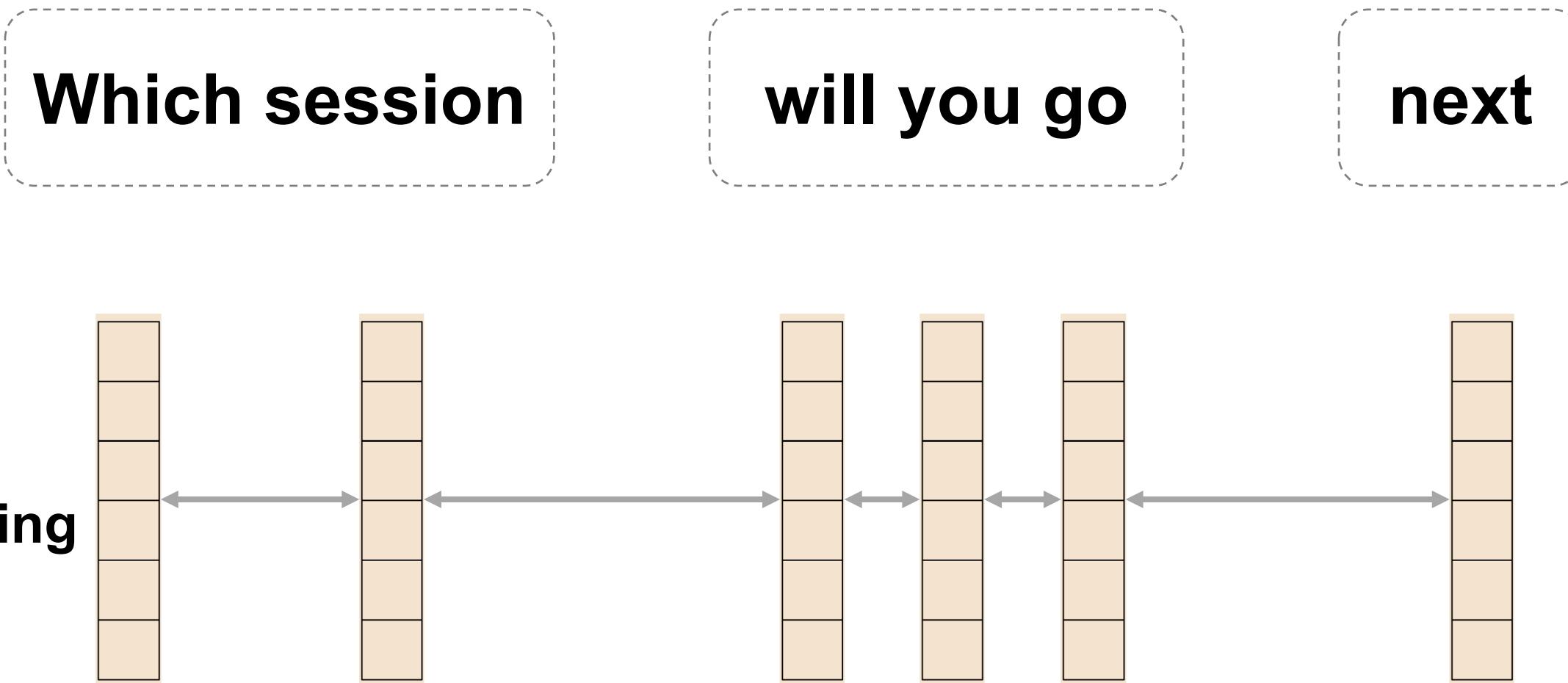
Estimate Event Representations

Which session

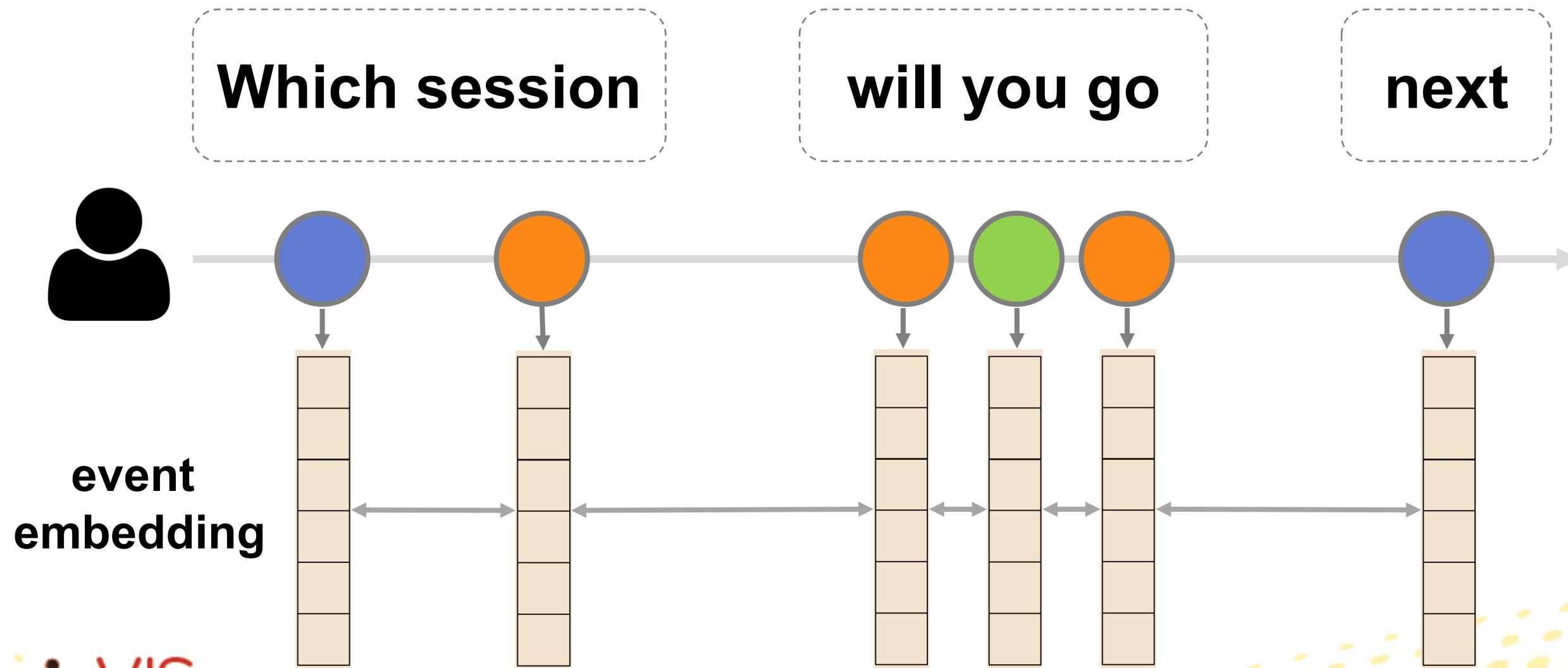
will you go

next

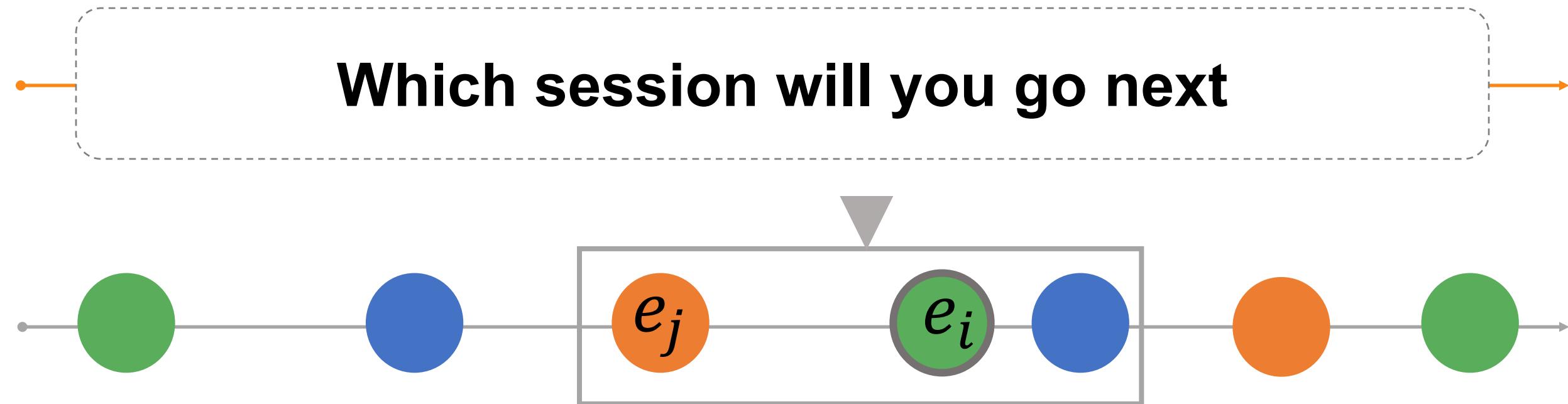
Estimate Event Representations



Estimate Event Representations



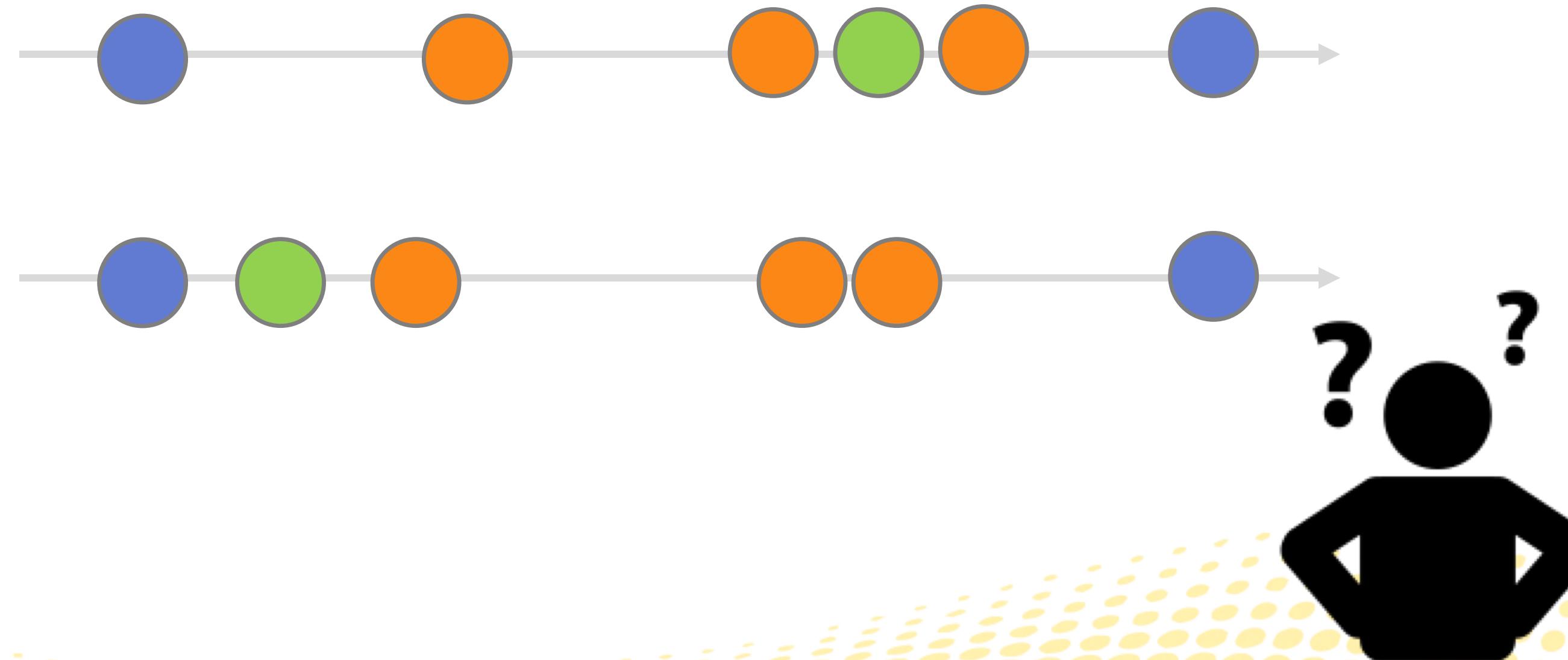
Estimate Event Representations



Training Sample: $(e_i, t_i), (e_j, t_j)$

$$\omega\langle(e_i, t_i), (e_j, t_j)\rangle = \begin{cases} \exp(-|t_i - t_j|/\theta) & |t_i - t_j| \leq T, \\ 0 & \text{otherwise} \end{cases}$$

Sequence Alignment



Sequence Alignment

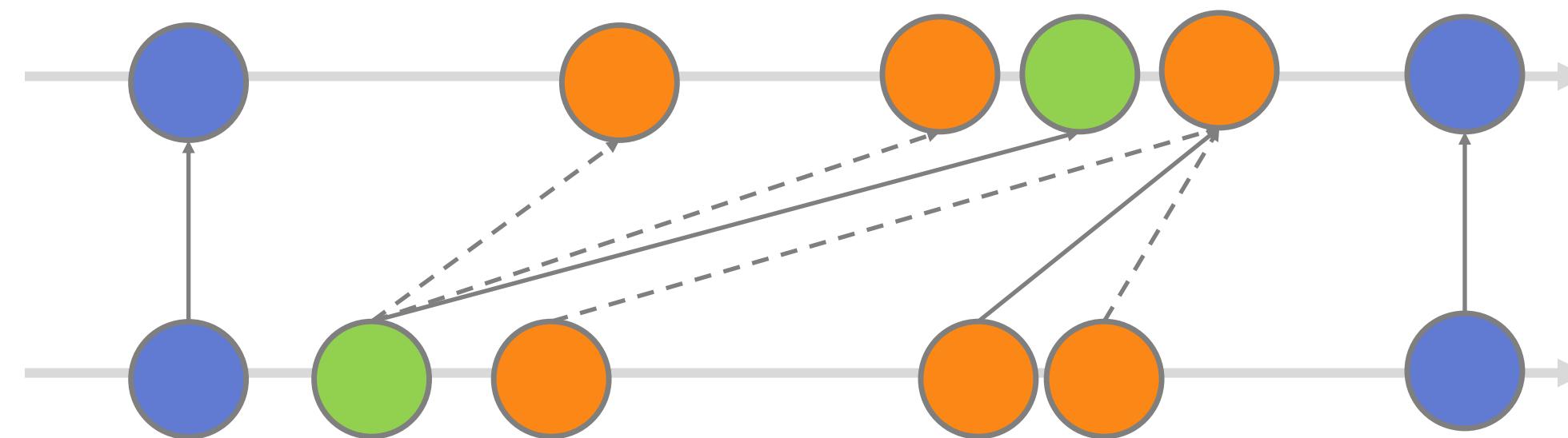


Time-based alignment



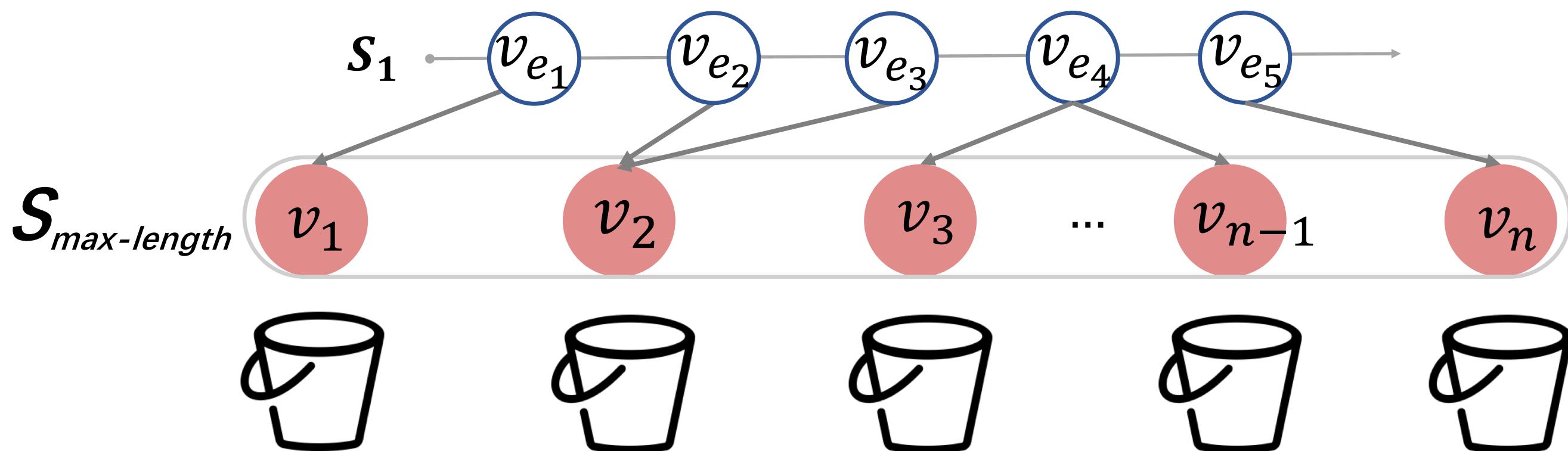
Order-based alignment

Sequence Alignment

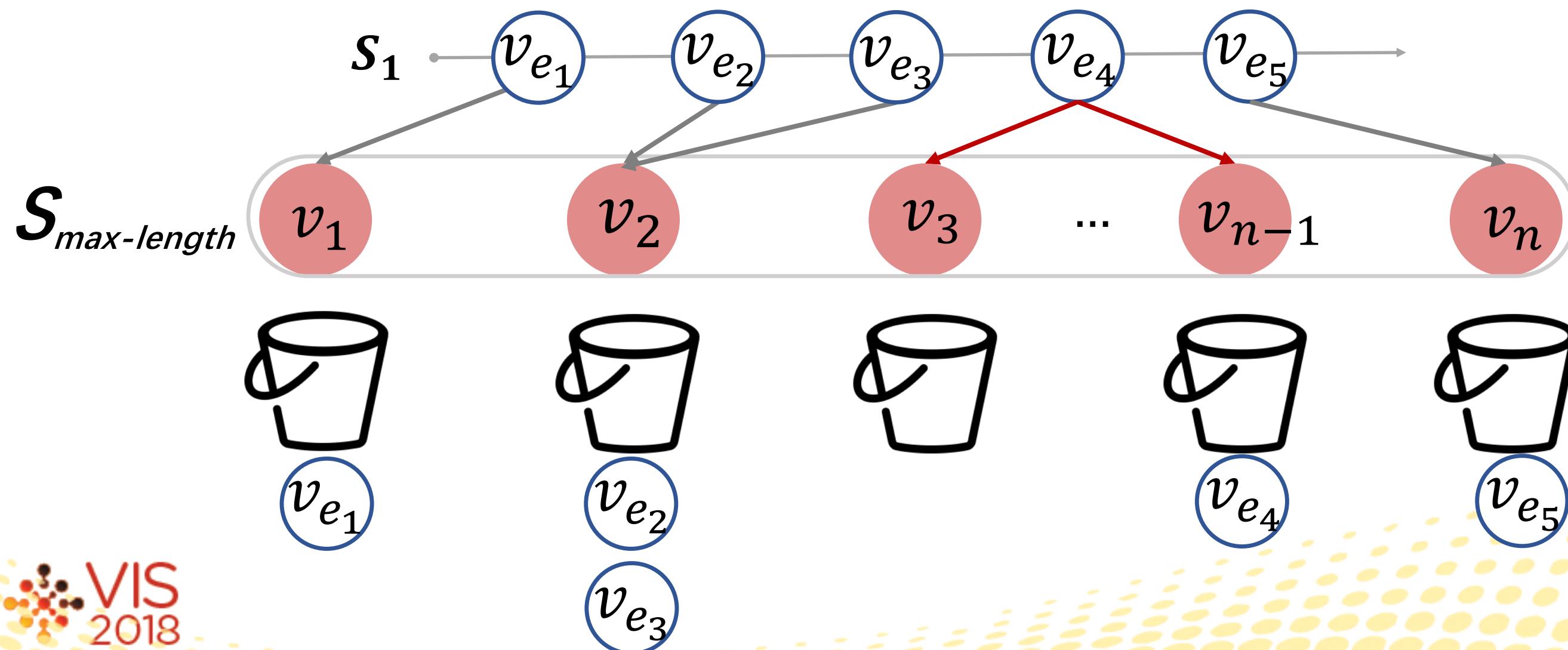


Dynamic Time Warping

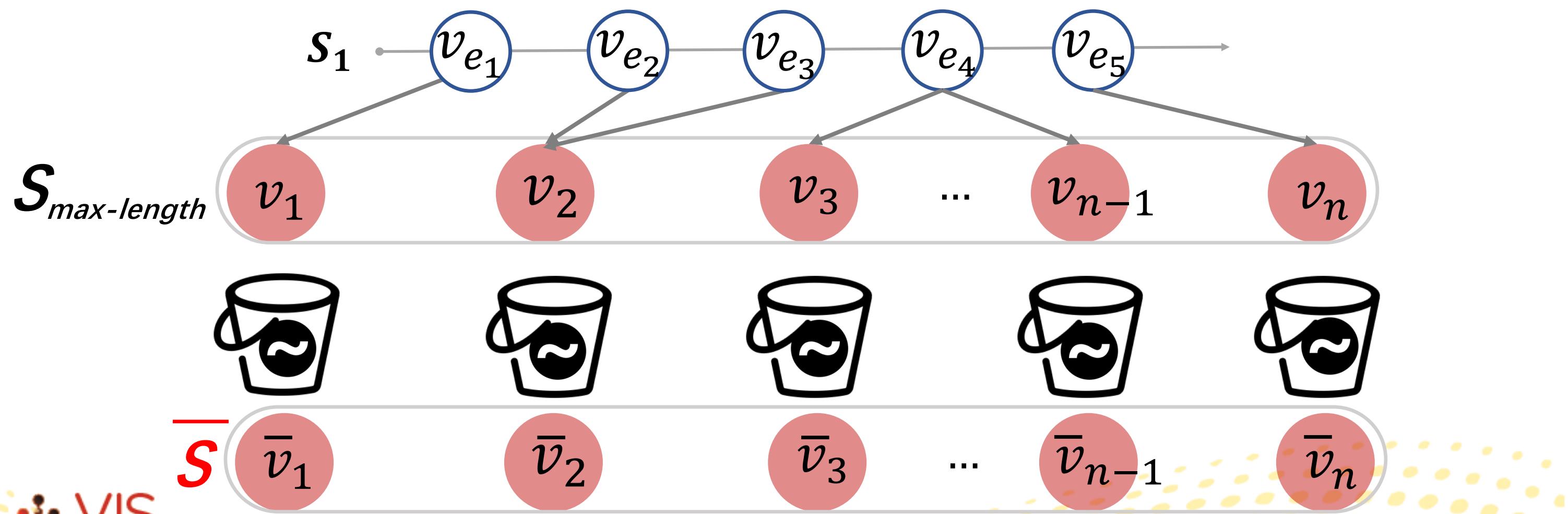
Sequence Alignment



Sequence Alignment



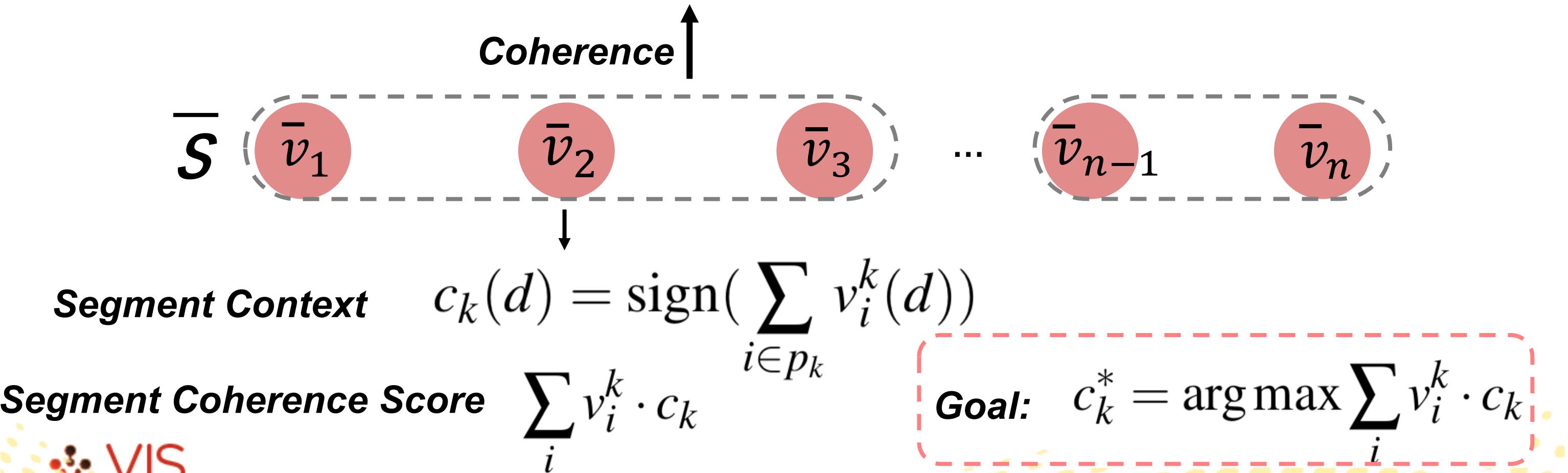
Sequence Alignment



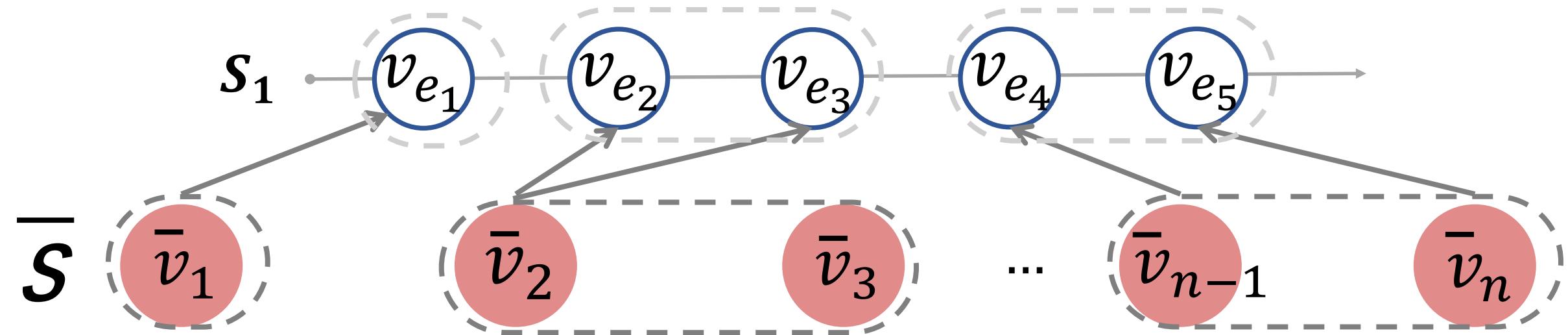
Sequence Segmentation



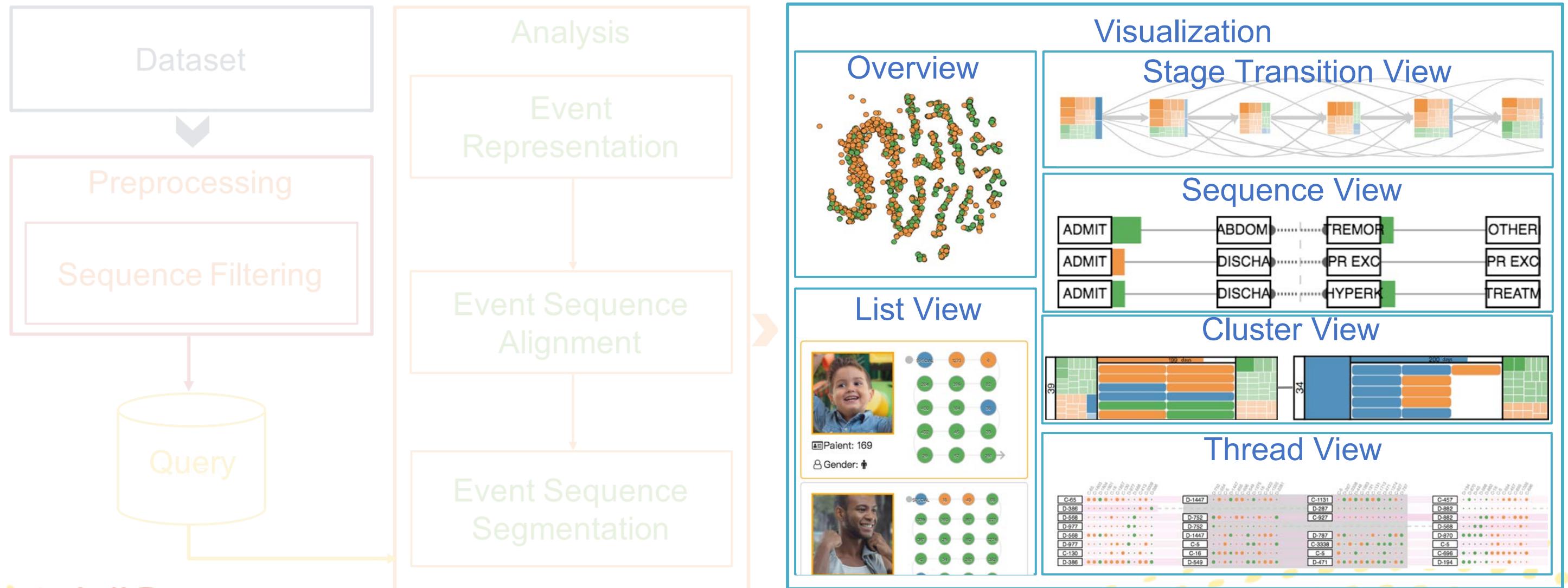
Sequence Segmentation



Sequence Segmentation

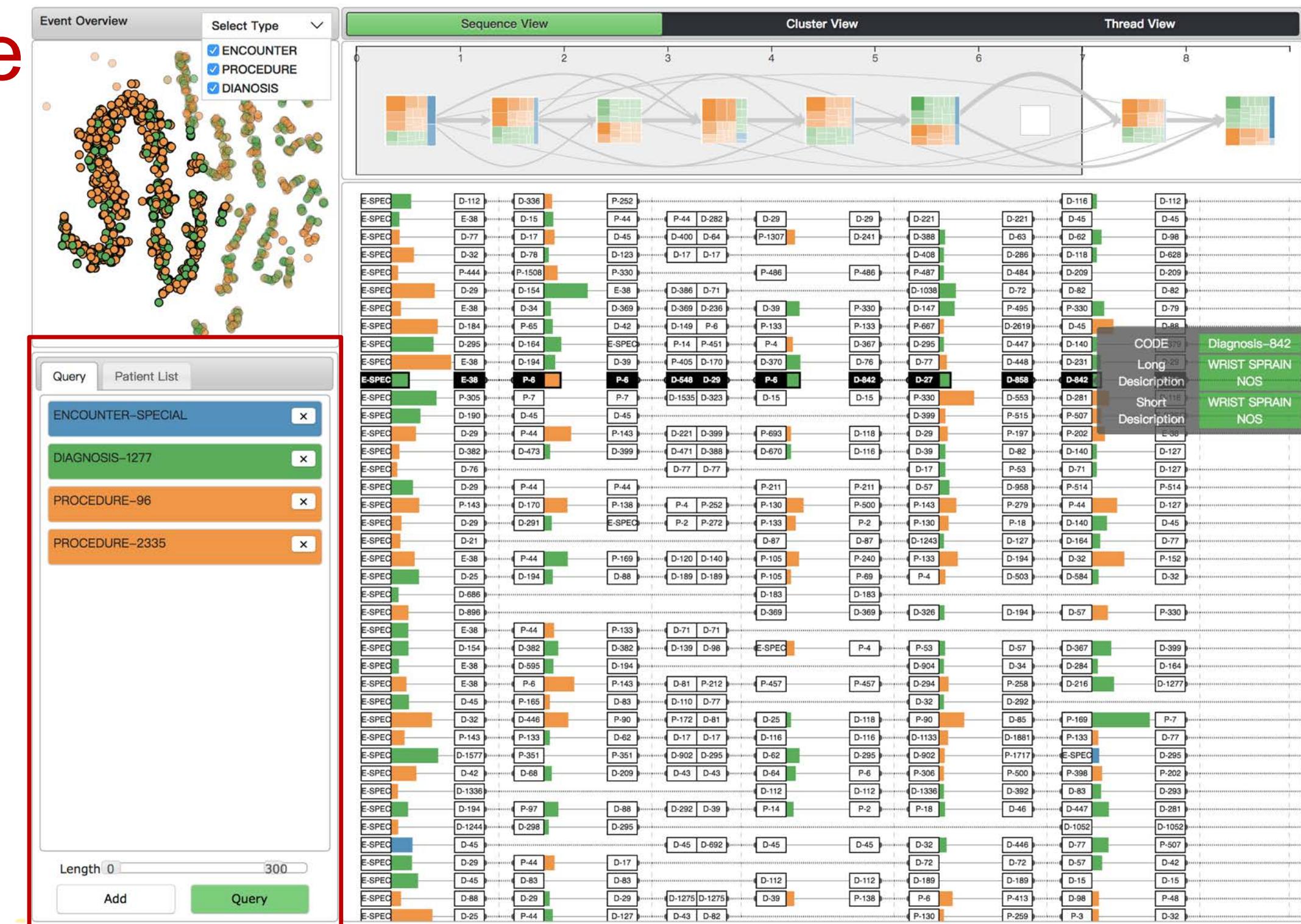
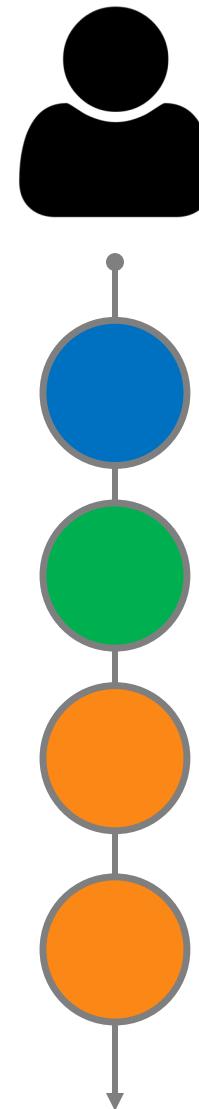


System Overview

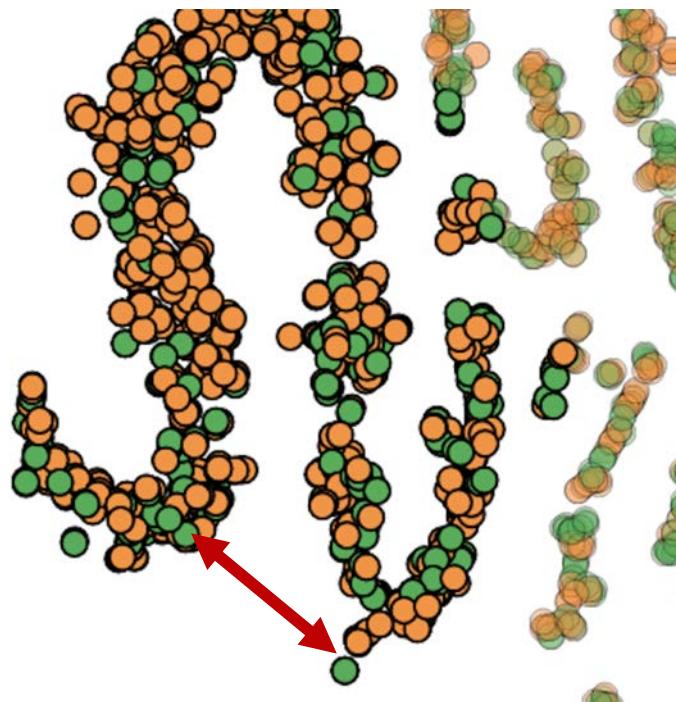


User Interface

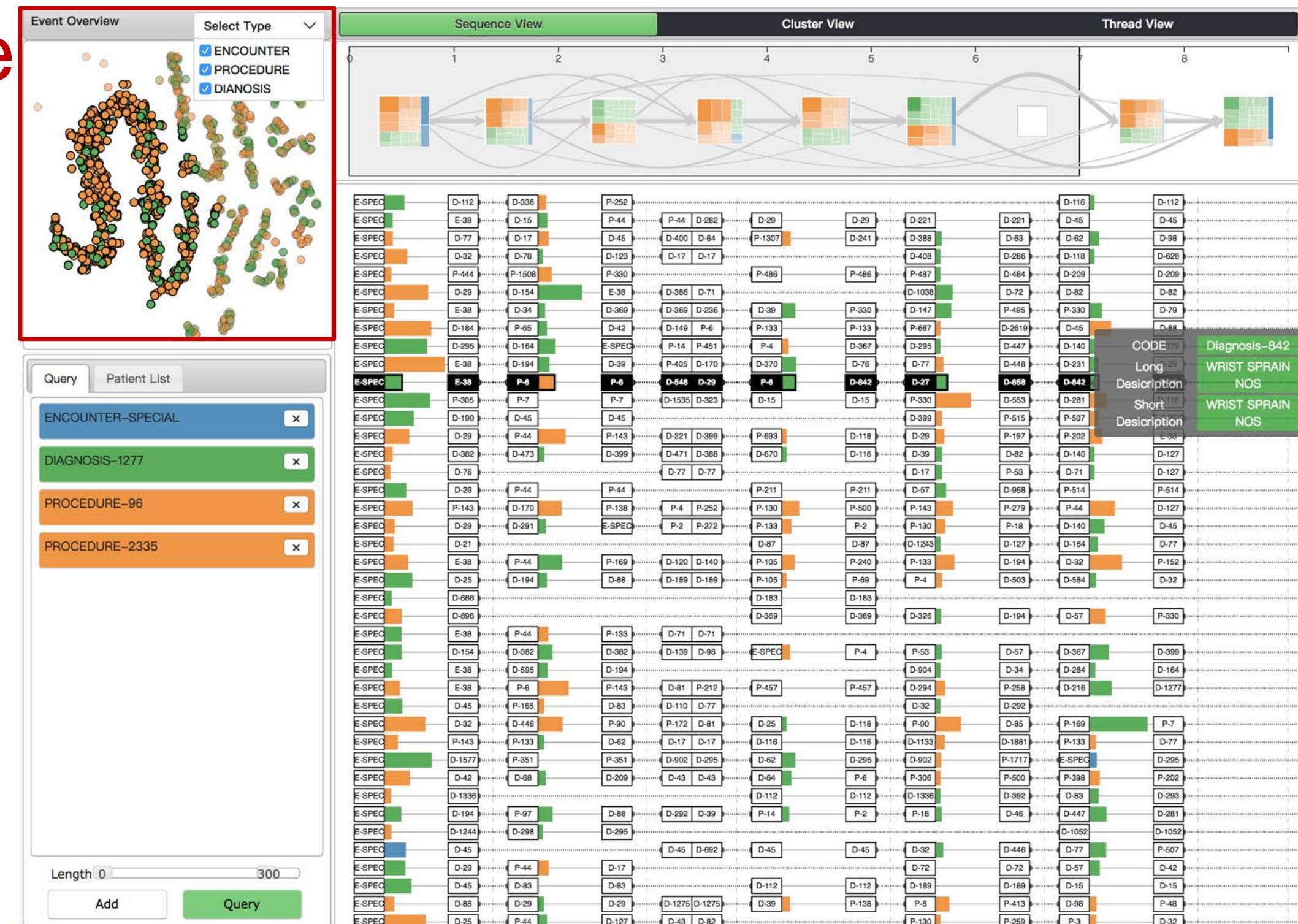
Milestone
Events



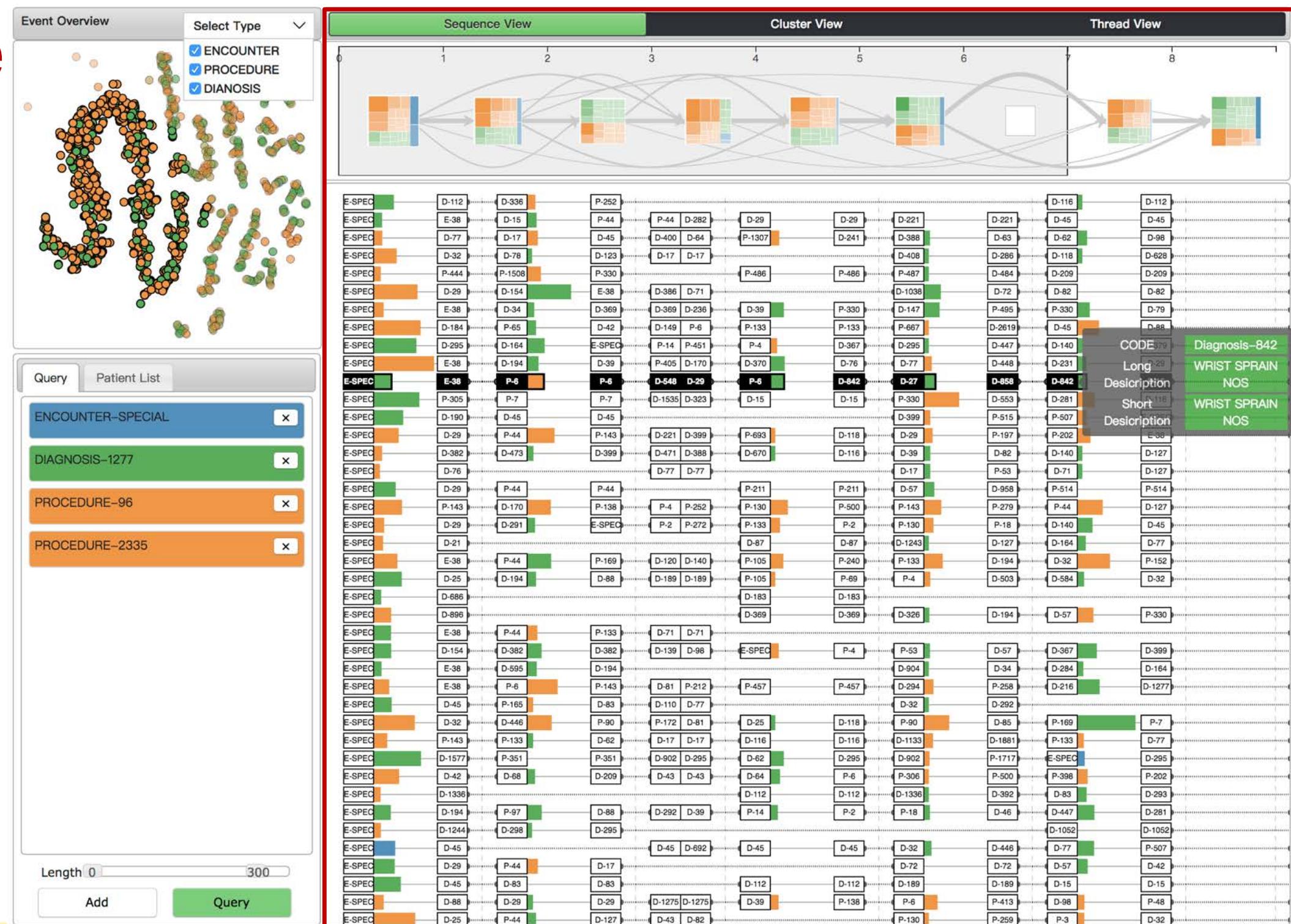
User Interface

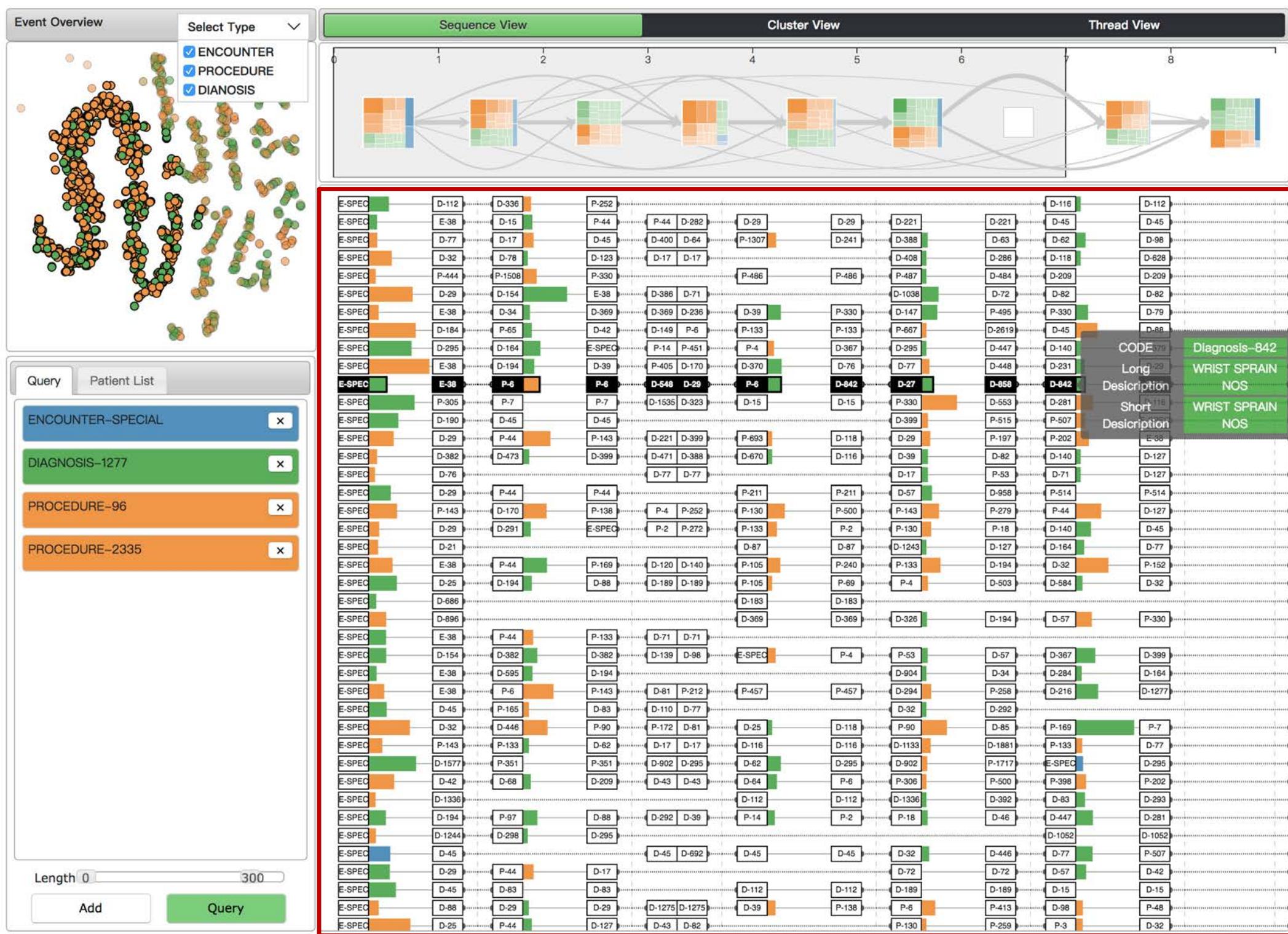


- Encounter
- Procedure
- Diagnosis

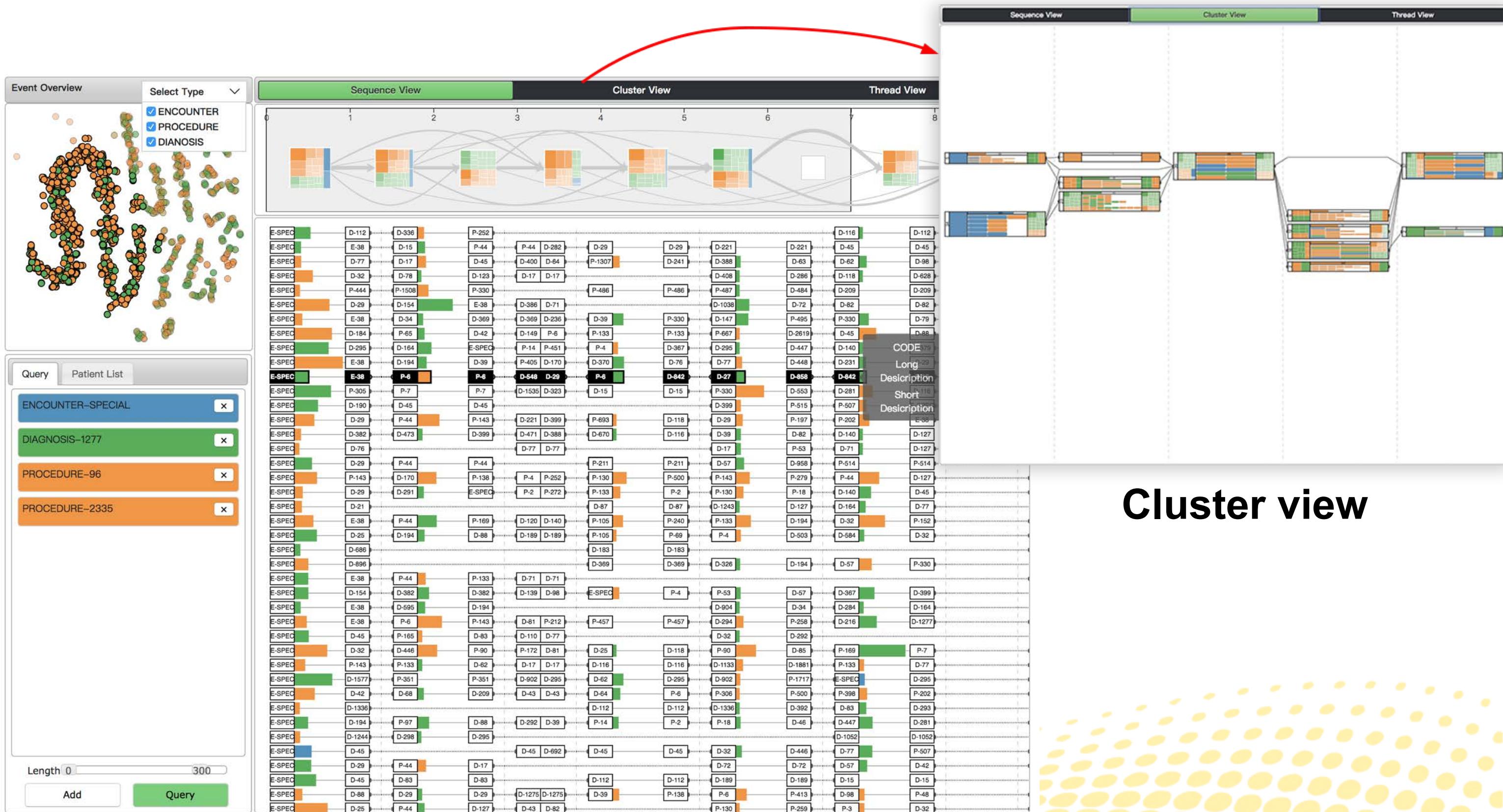


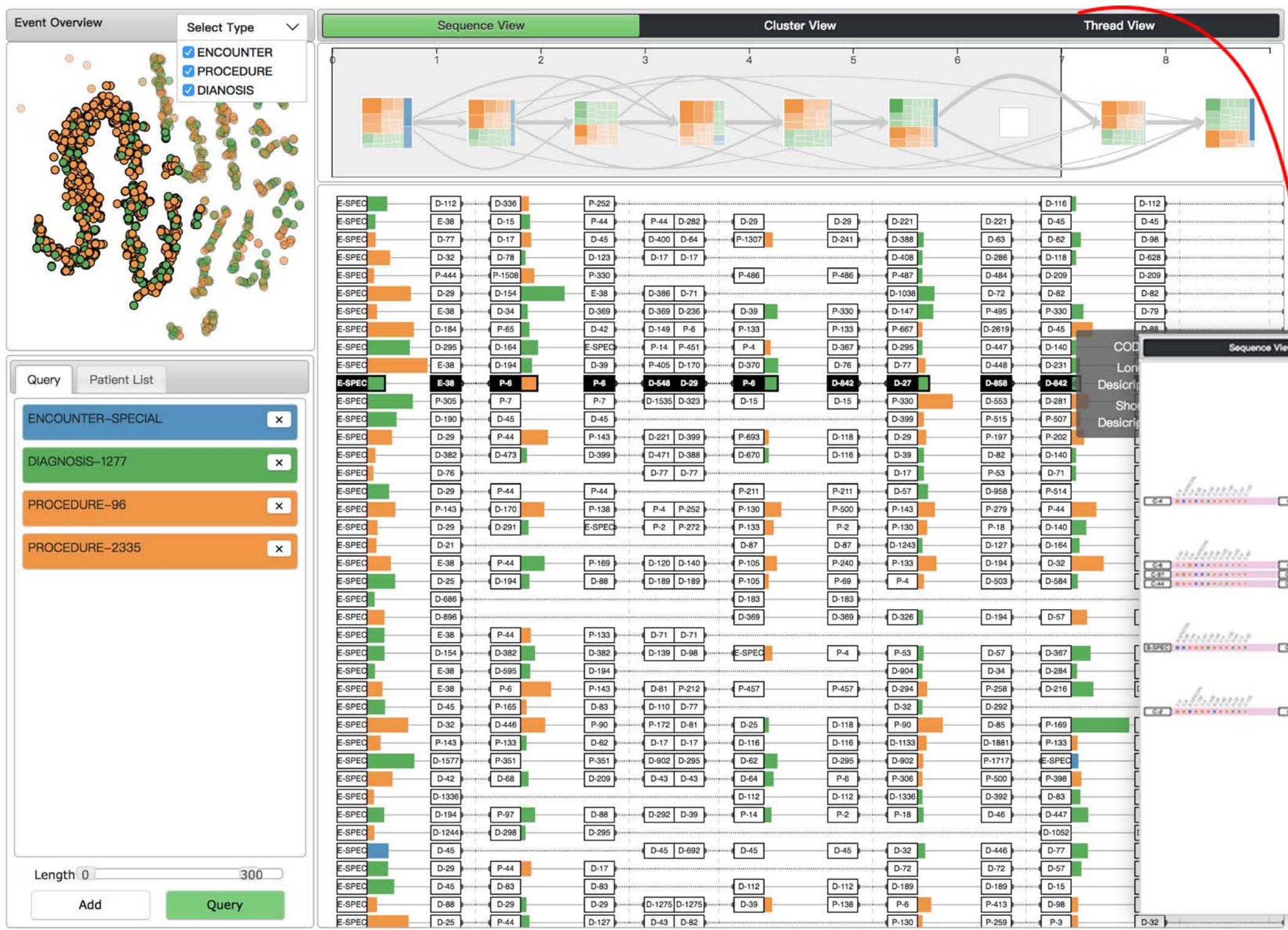
User Interface



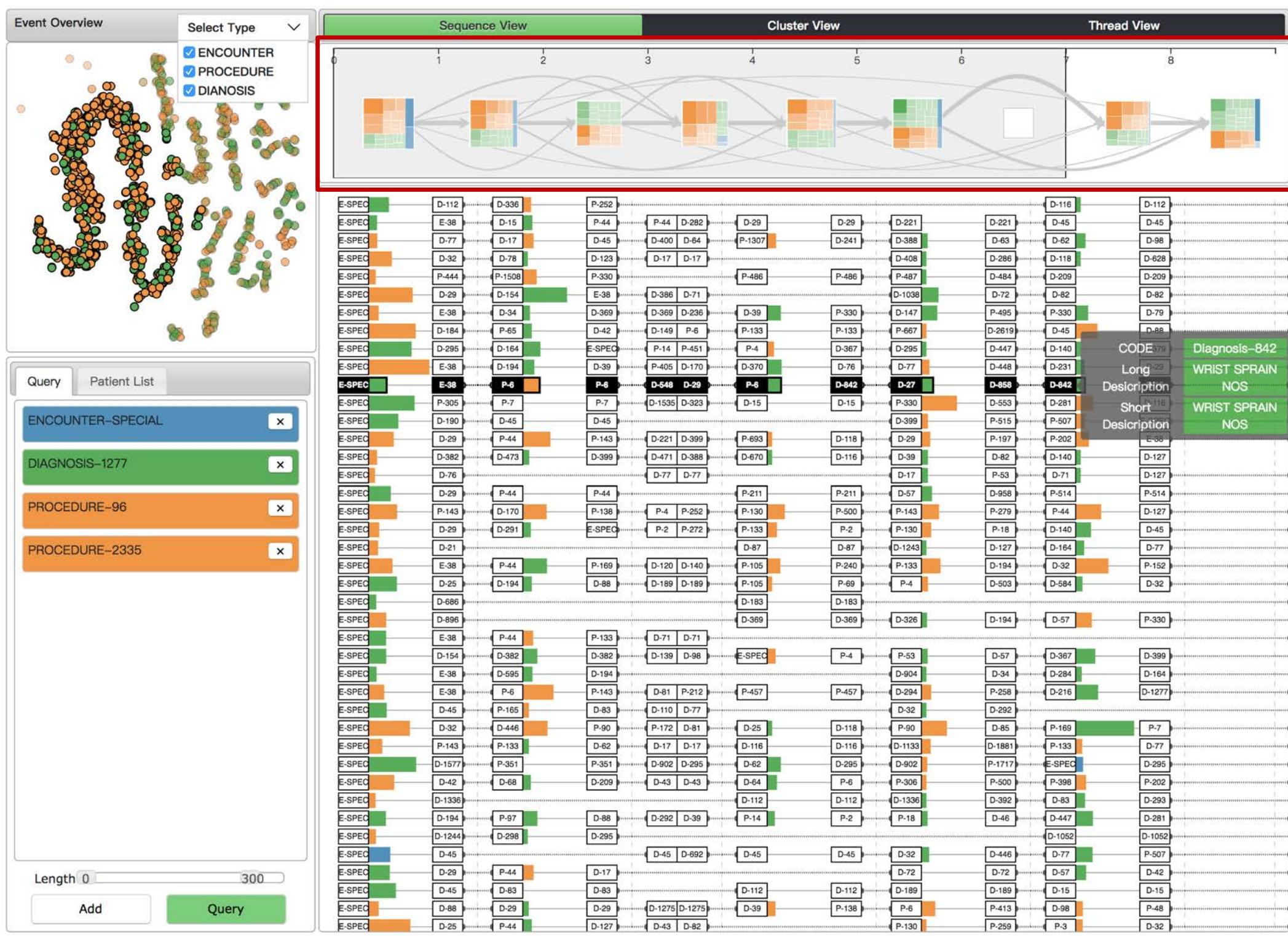


Sequence view

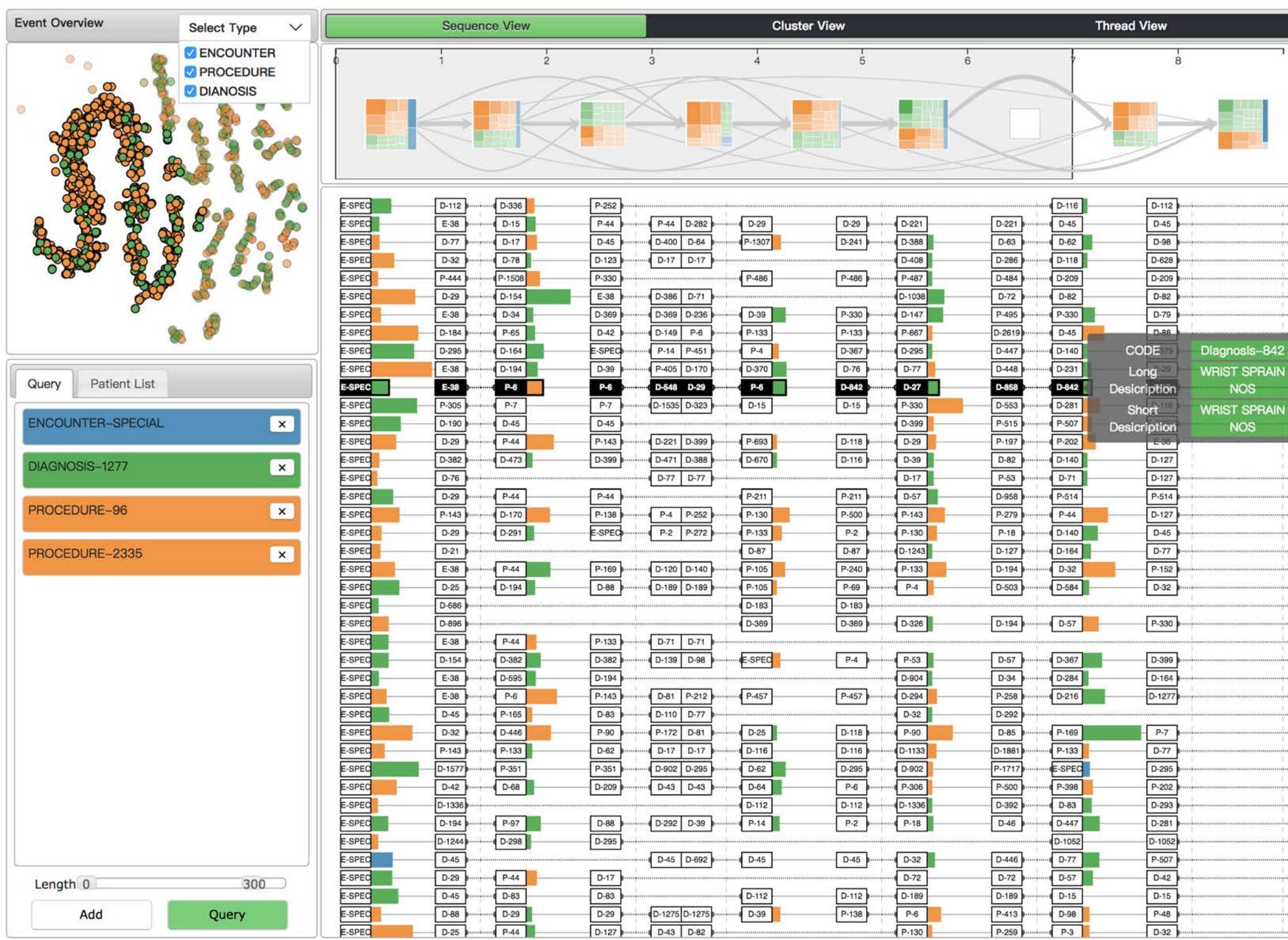




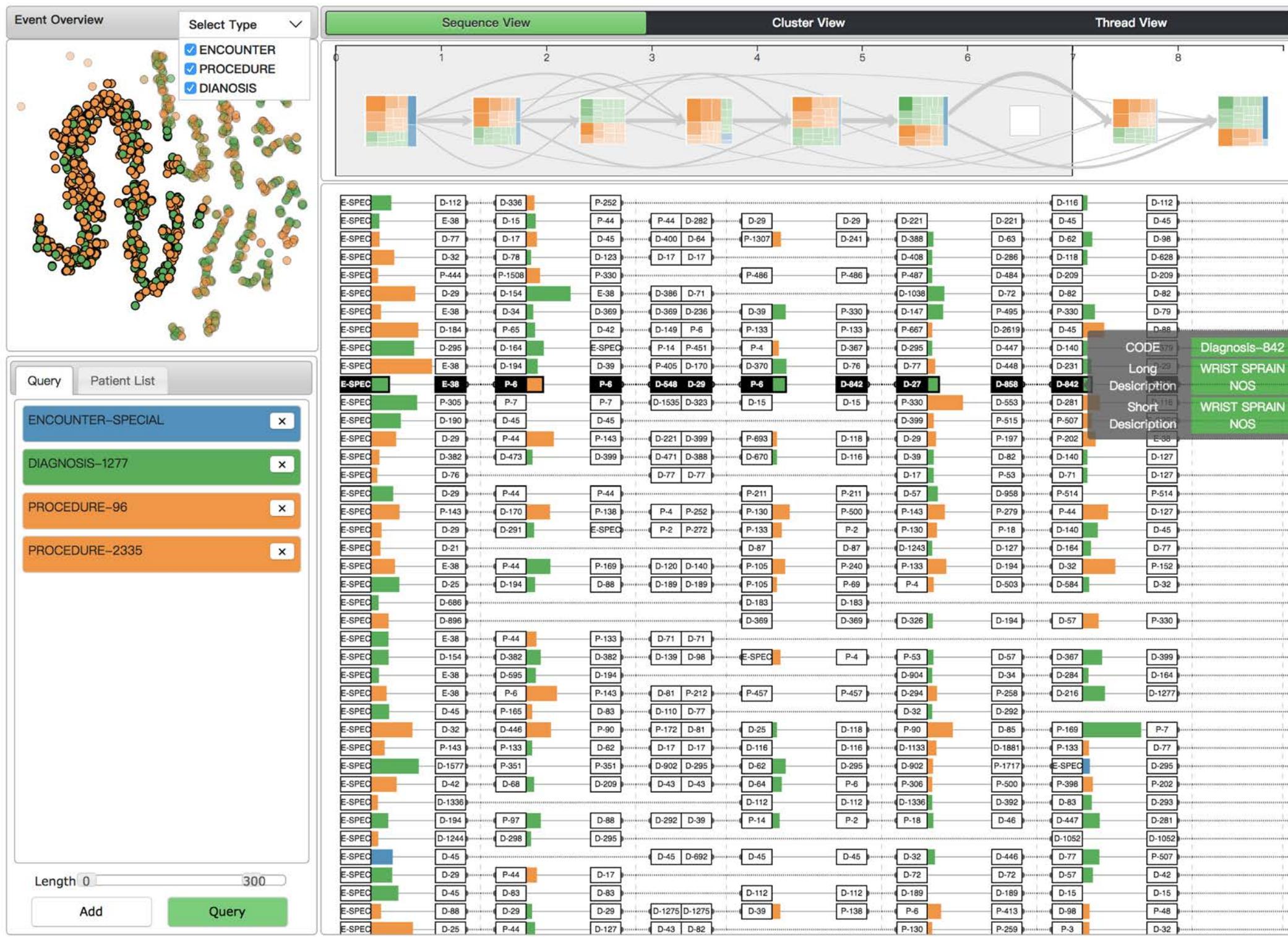
Thread view



Stage transition view

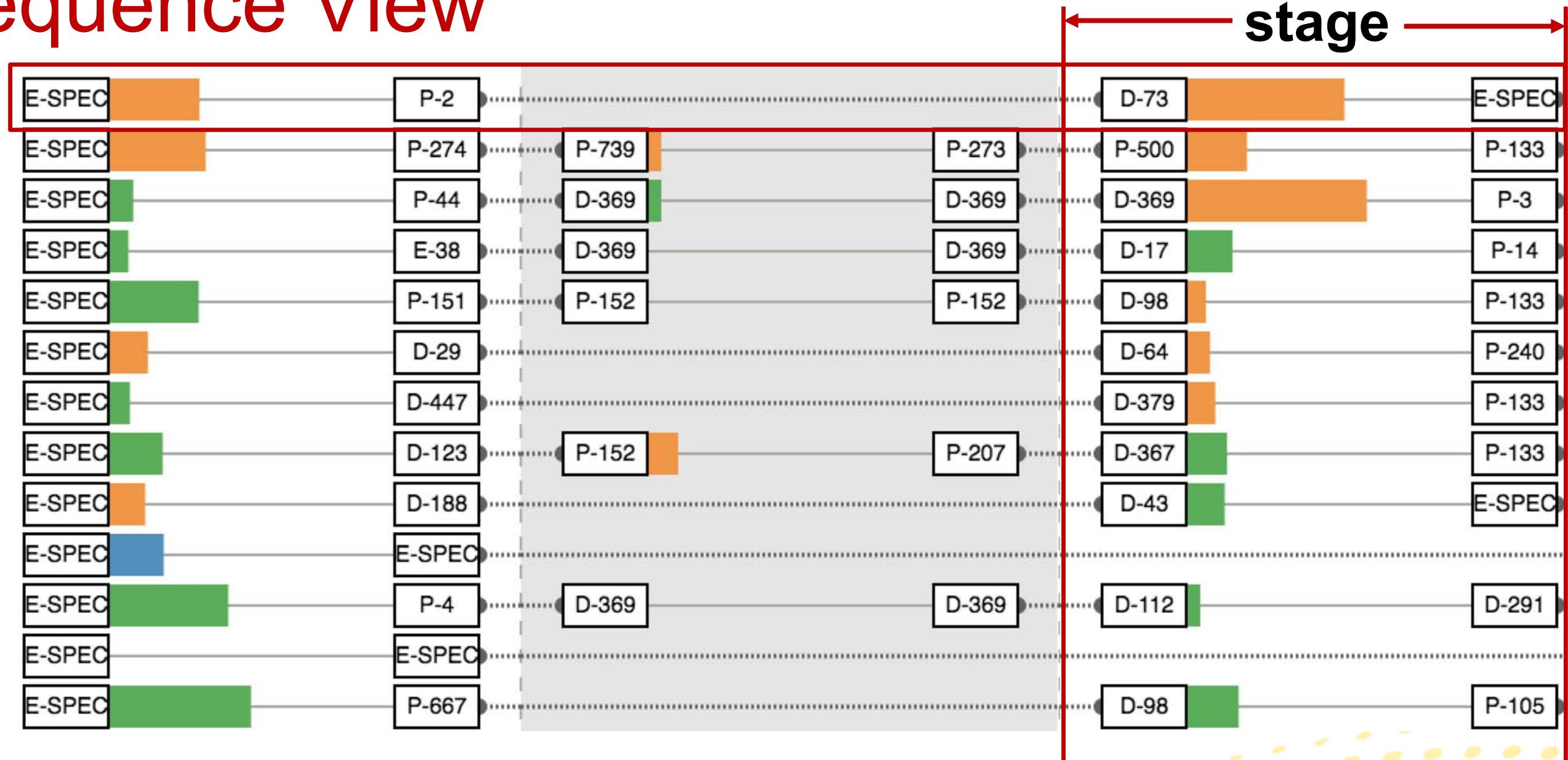


- Summarization Level
- Sequence view
- Cluster view
- Stage transition view

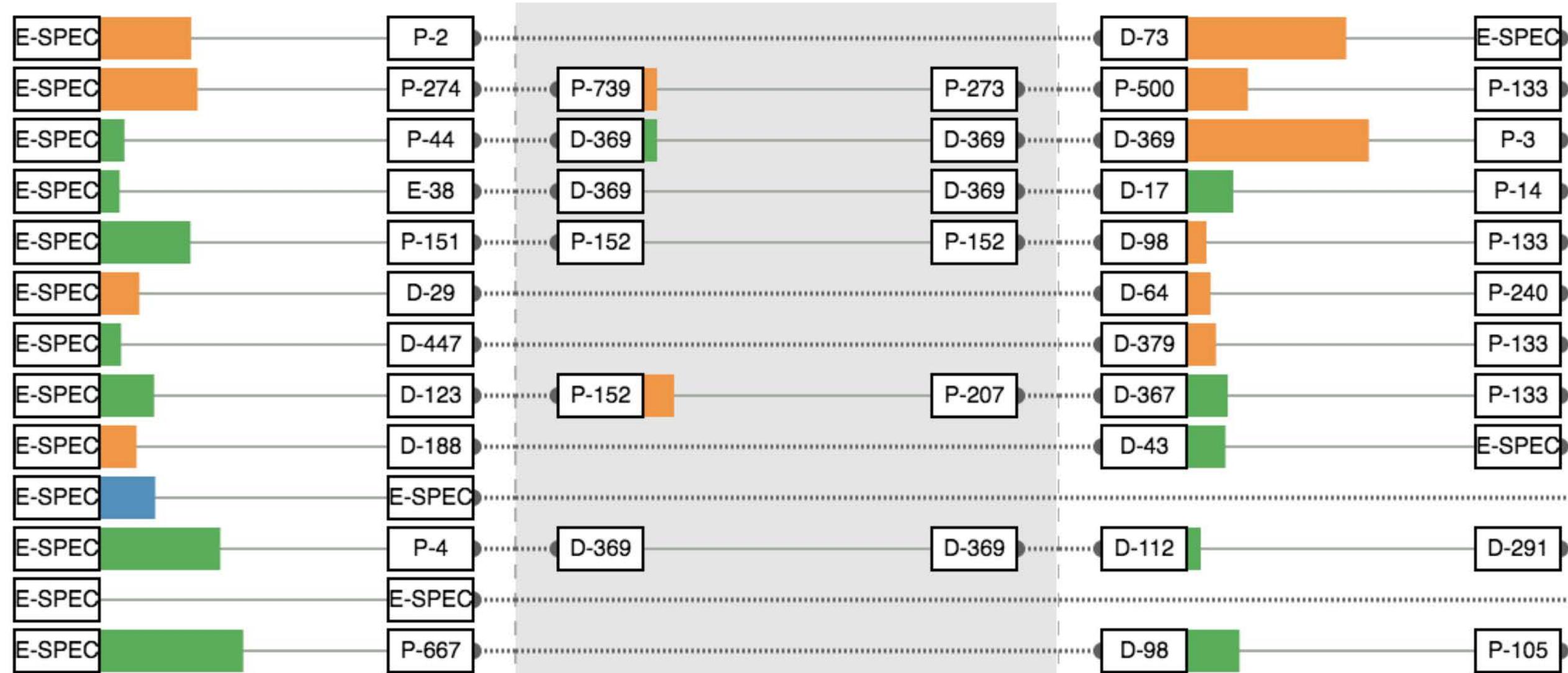


- Summarization Level
- ✓ **Sequence view**
- **Cluster view**
- **Stage transition view**

Sequence View



Sequence View



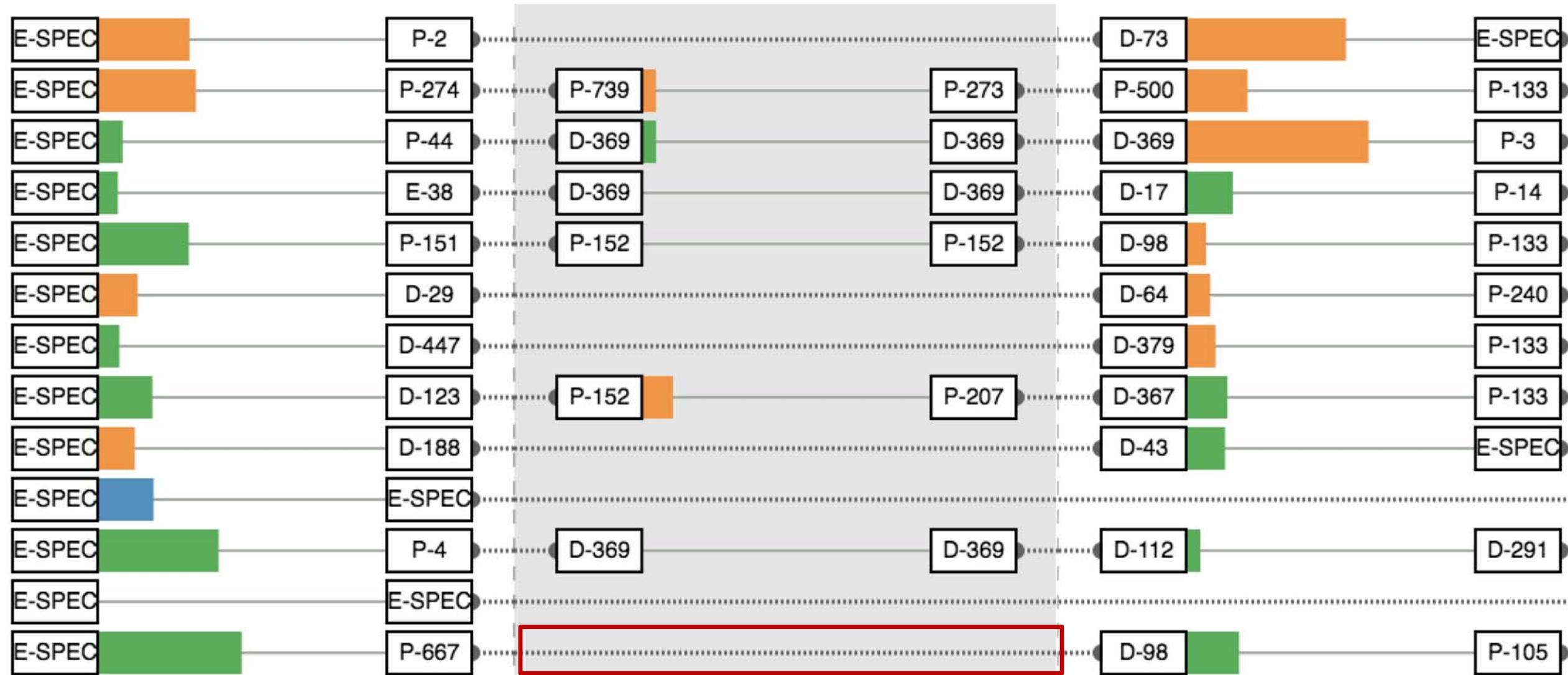
stage segment

D-98



P-105

Sequence View

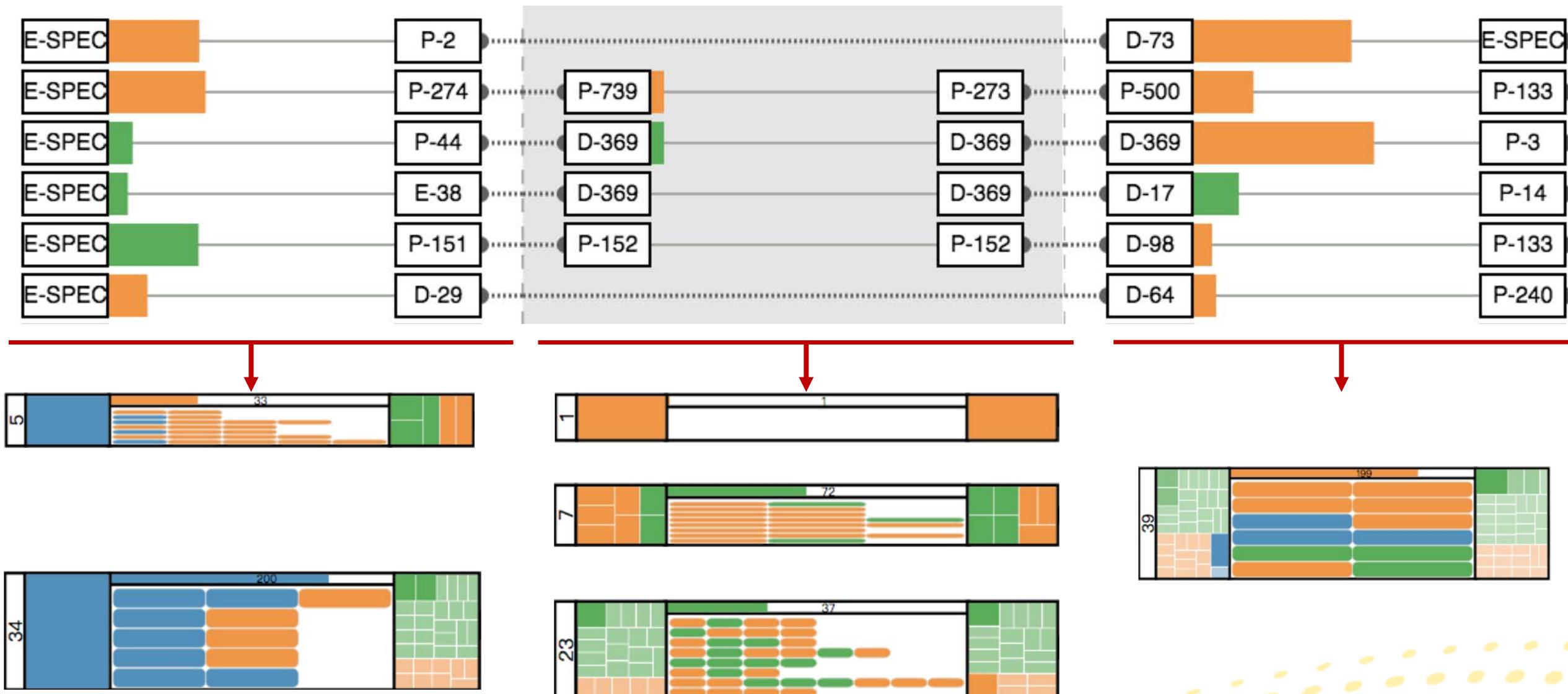


stage segment

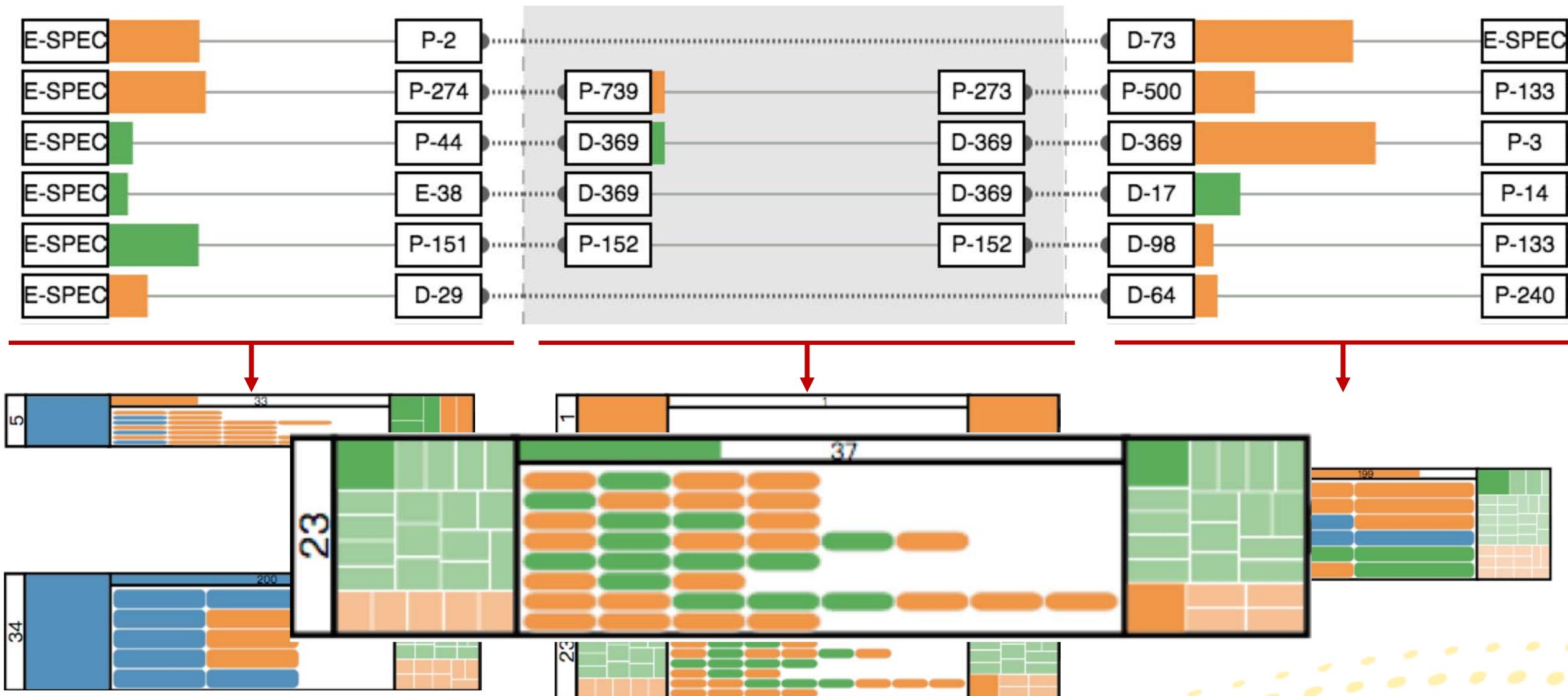
D-98

P-105

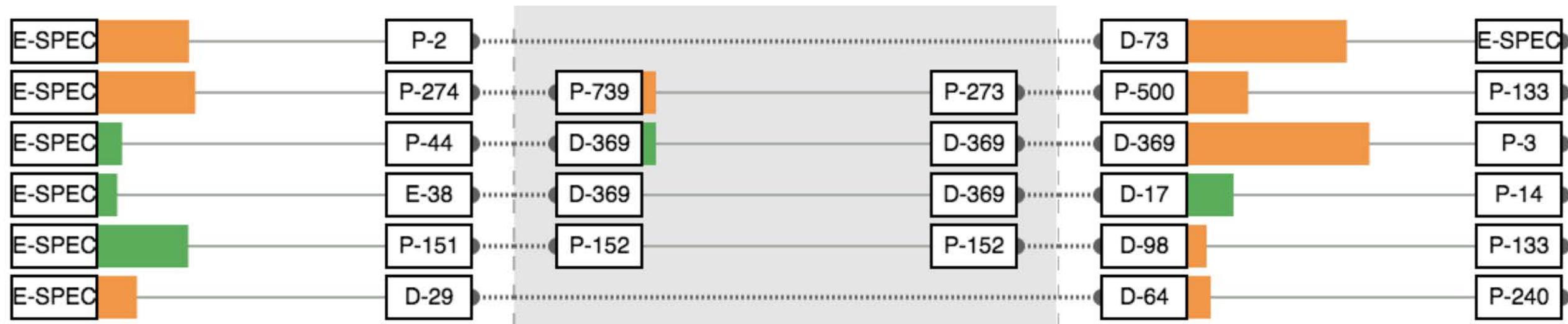
Cluster View



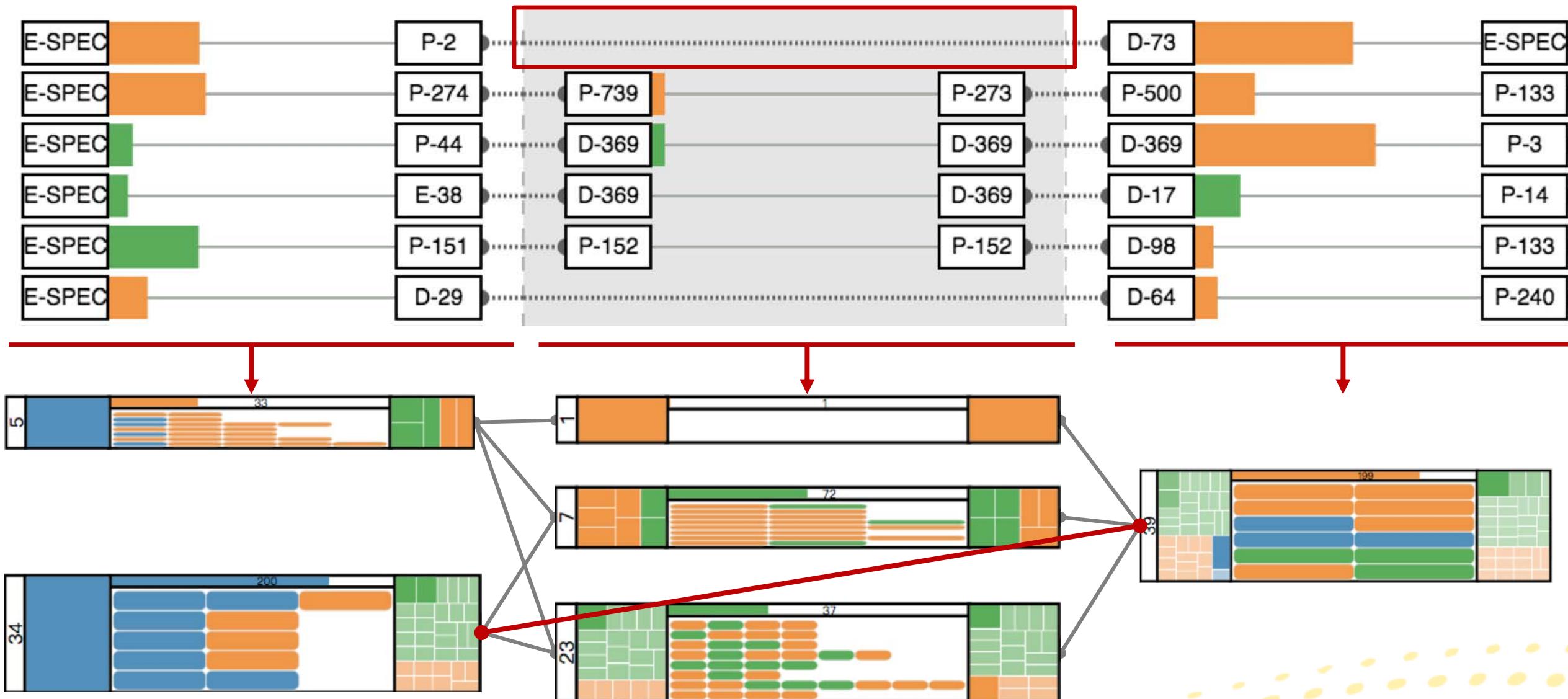
Cluster View



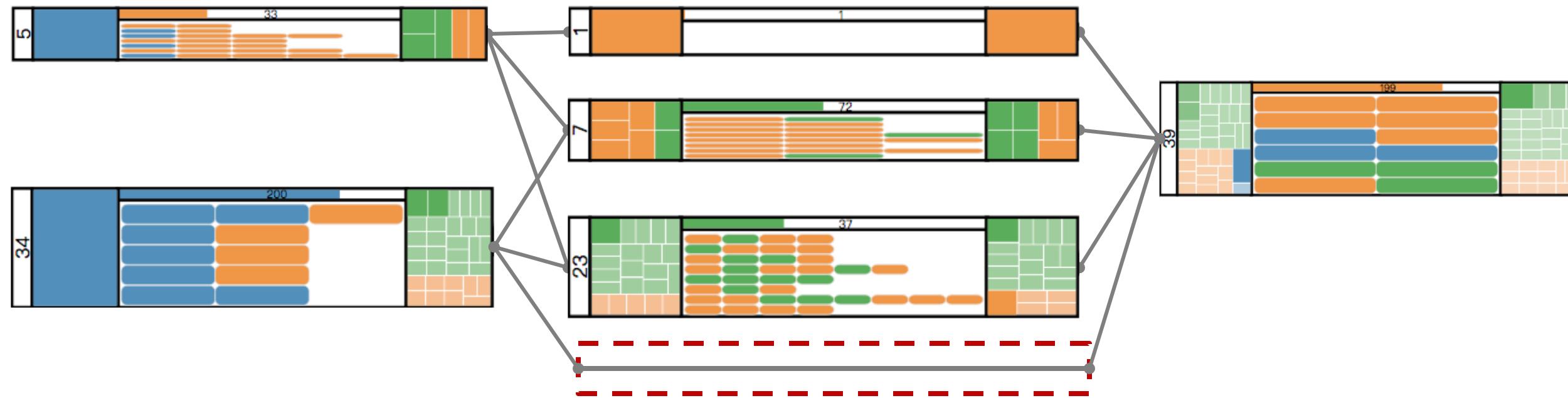
Cluster View



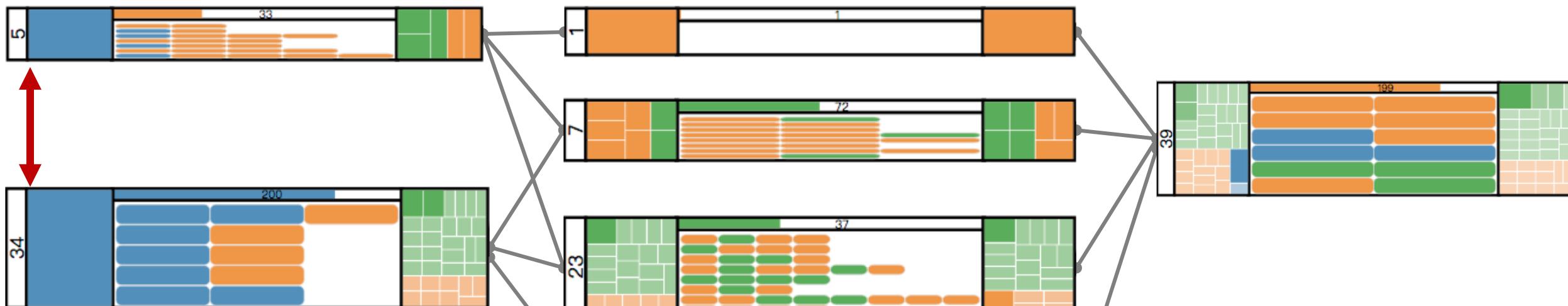
Layout Algorithm of Cluster View



Step#1 Insertion of the Intermediate Nodes



Step#2 Positioning of the Cluster Nodes

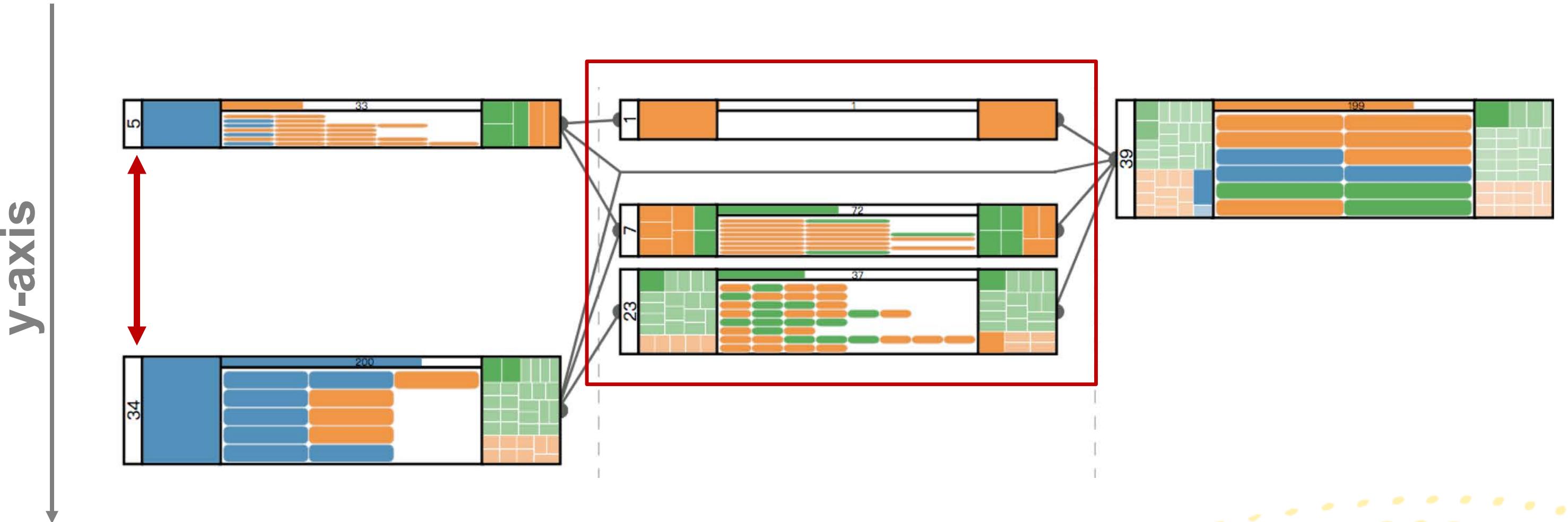


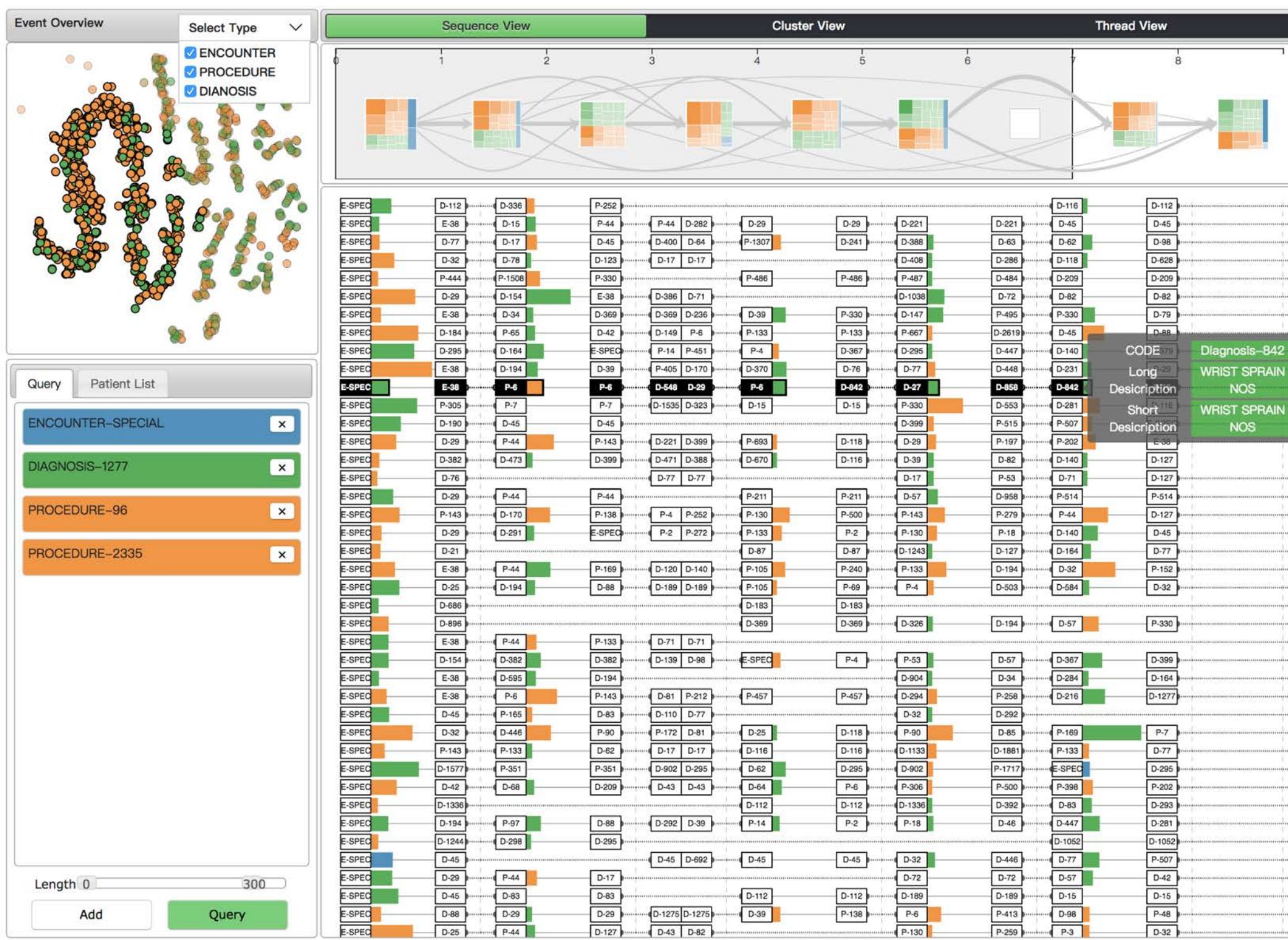
$$\sum_{t=0}^T \left(\alpha \sum_{i < j} \omega_{ij}^t \|y_i^t - y_j^t\|^2 + (1 - \alpha) \sum_i \sum_k \|y_i^t - \theta_{ik} \cdot y_{ik}^{t-1}\|^2 \right)$$

Group similar nodes

Minimize link crossing

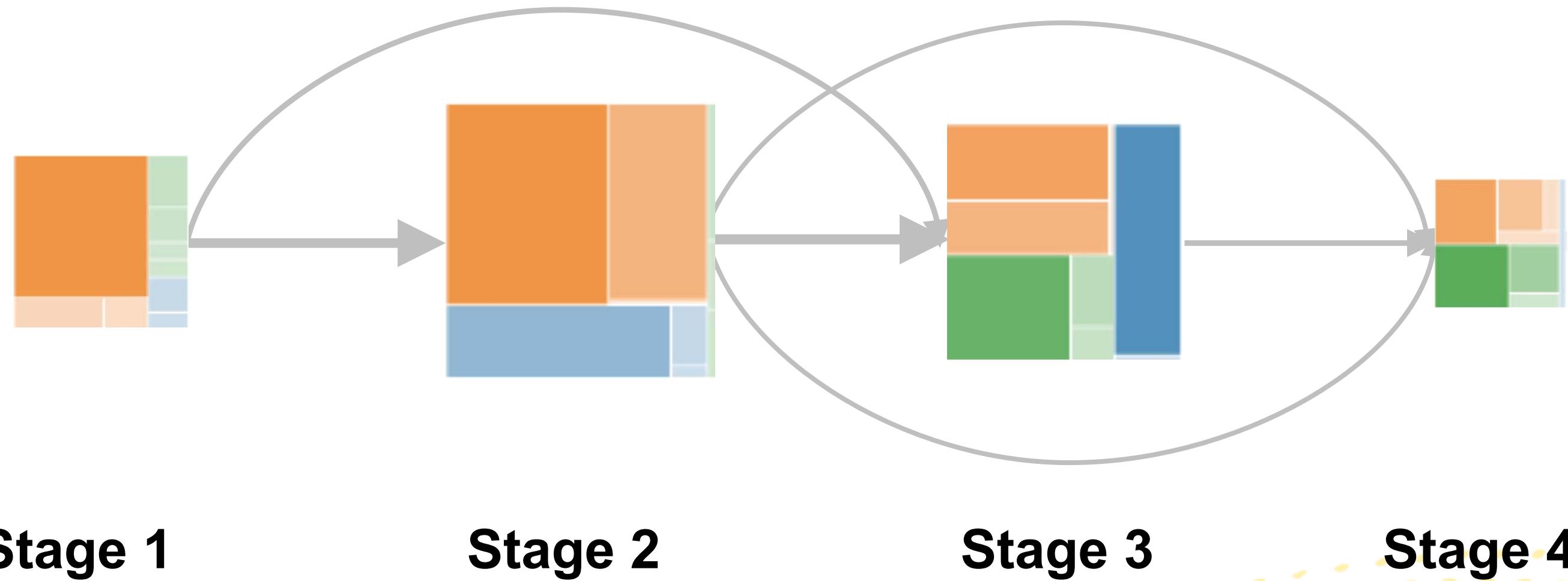
Step#2 Positioning of the Cluster Nodes





- Summarization Level
- ✓ Sequence view
- ✓ Cluster view
- ✓ Stage transition view

Stage Transition View



Stage 1

Stage 2

Stage 3

Stage 4

Video Demo

Event Overview Select Type ▾ Sequence View Cluster View Thread View

Query Entity List

- ENCOUNTER-SPECIAL
- PROCEDURE-99
- DIAGNOSIS-1277
- PROCEDURE-667
- DIAGNOSIS-504

Length

Add Query

Keyword:

DIAGNOSIS-504

DIAGNOSIS-509

DIAGNOSIS-508

Add

VIS
2018

Evaluation

- Case studies on **MIMIC** and **academic behavior** datasets
- Expert interview
- Performance evaluation on **algorithm** and **visualization**

Evaluation

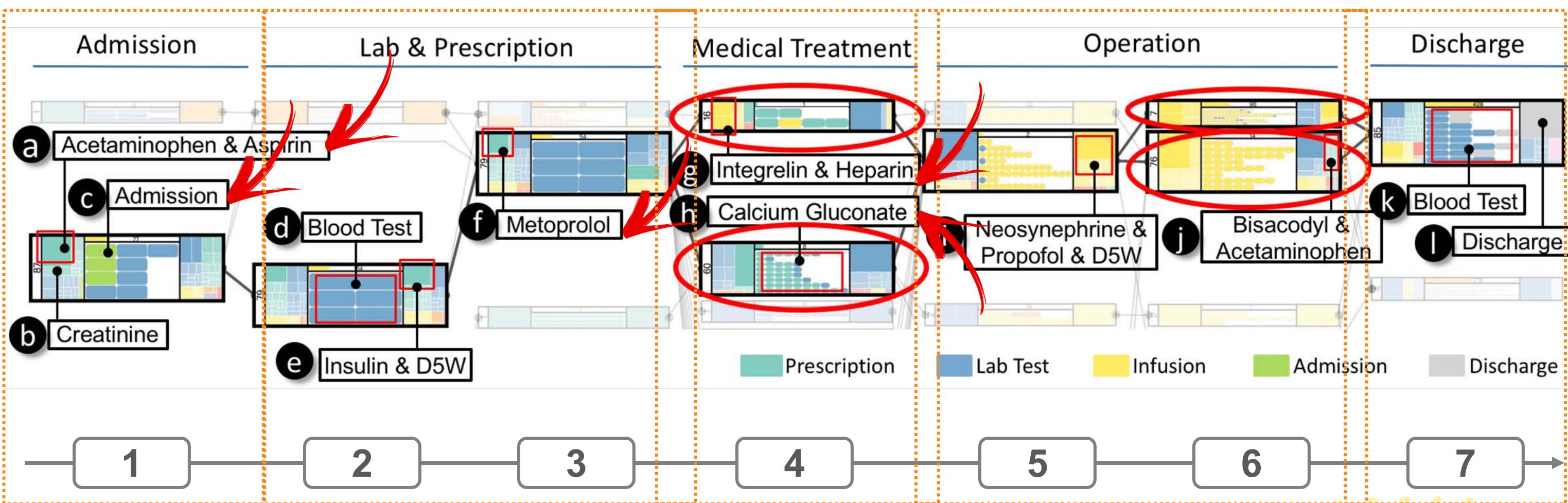
- Case studies on **MIMIC** and academic behavior datasets
- Expert interview
- Performance evaluation on **algorithm** and **visualization**

Case Study – MIMIC Dataset

- Electronic health records of 46,521 patients who were admitted to the intensive care unit (ICU)
- 11,000+ event types, organized into seven categories:
 - hospital admission and discharge
 - ICU admission and discharge
 - Death
 - Prescriptions
 - Infusions
 - Laboratory tests
 - Microbiological test.

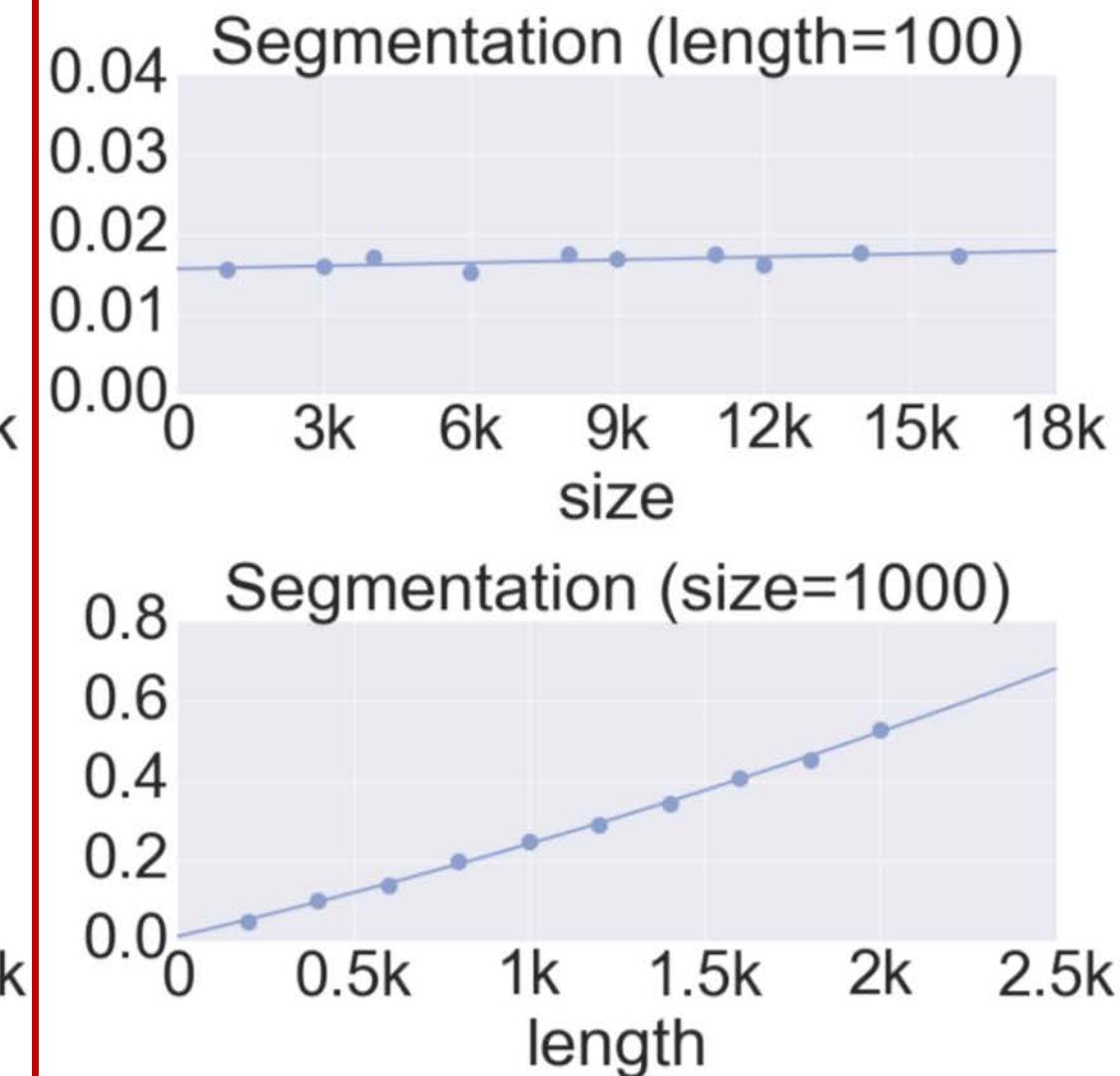
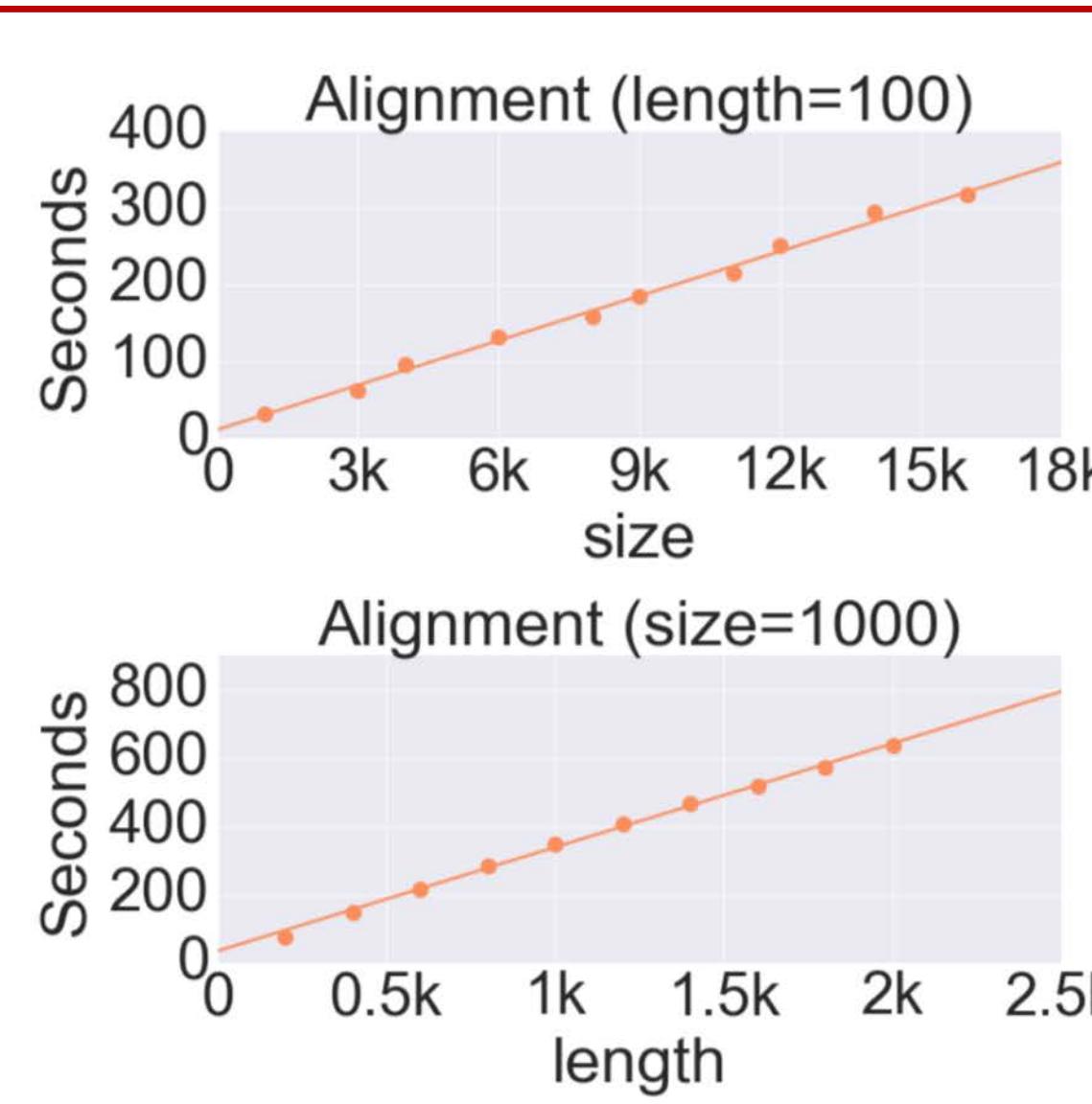


Case Study – MIMIC Dataset



Performance Evaluation

- Scalability of the progression analysis algorithm



Conclusion & Future Directions

- ✓ An unsupervised progression analysis algorithm for event sequence data
- ✓ An interactive visual analytics system for result interpretation and exploration



Improve the scalability of the stage analysis algorithm



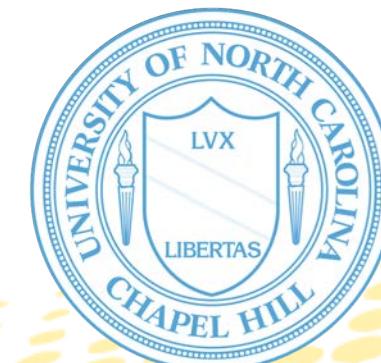
Formal usability experiment



Thank you!



Shunan Guo¹, Zhuochen Jin², David Gotz³, Fan Du⁴, Hongyuan Zha¹, Nan Cao² ☐



Q & A

