



Visual Progression Analysis of Event Sequence Data

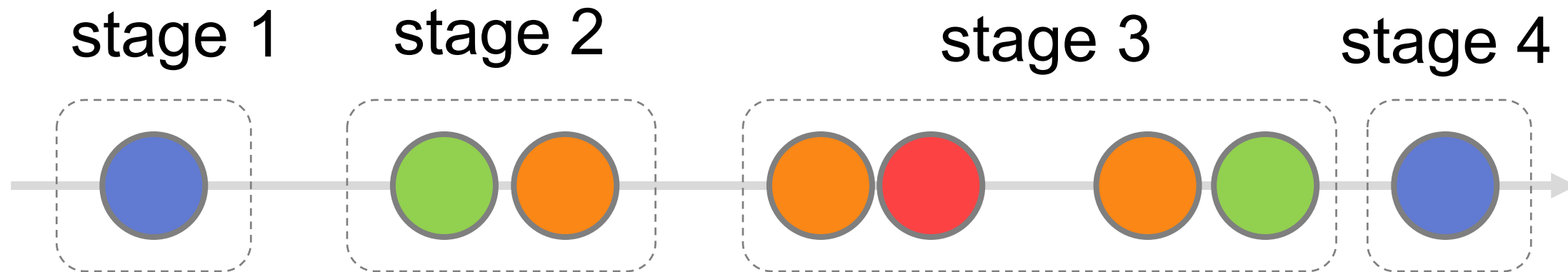
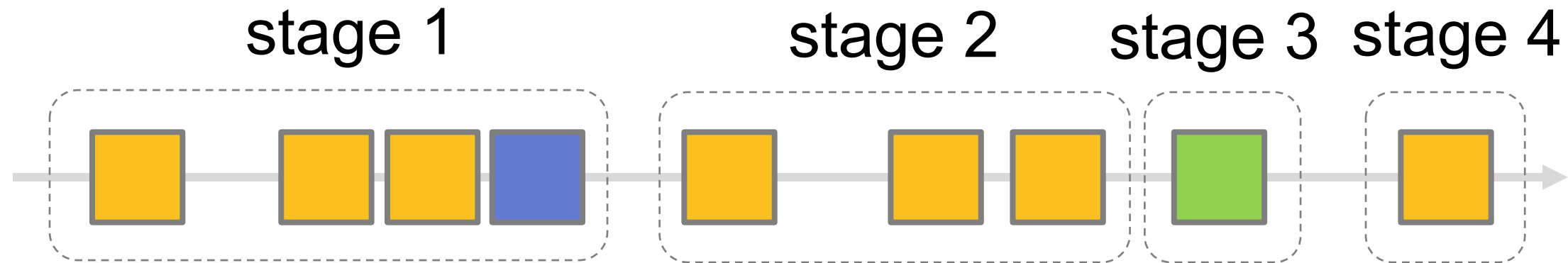
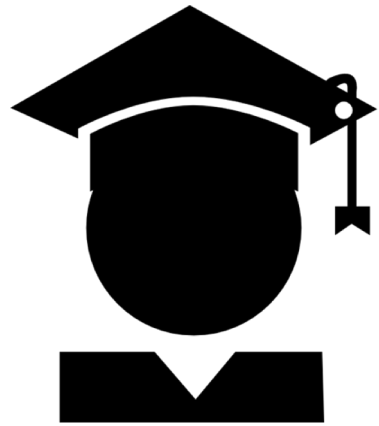
Shunan Guo¹, Zhuochen Jin², David Gotz³, Fan Du⁴, Hongyuan Zha¹, Nan Cao² ✉



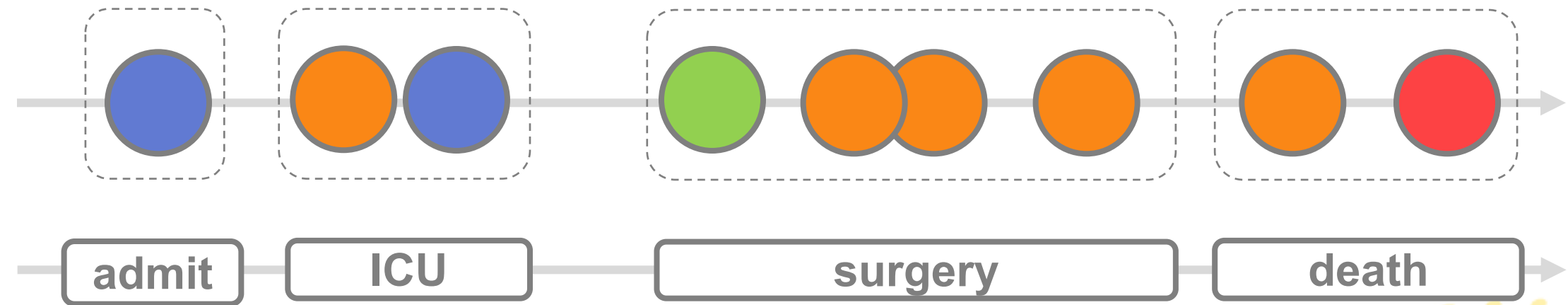
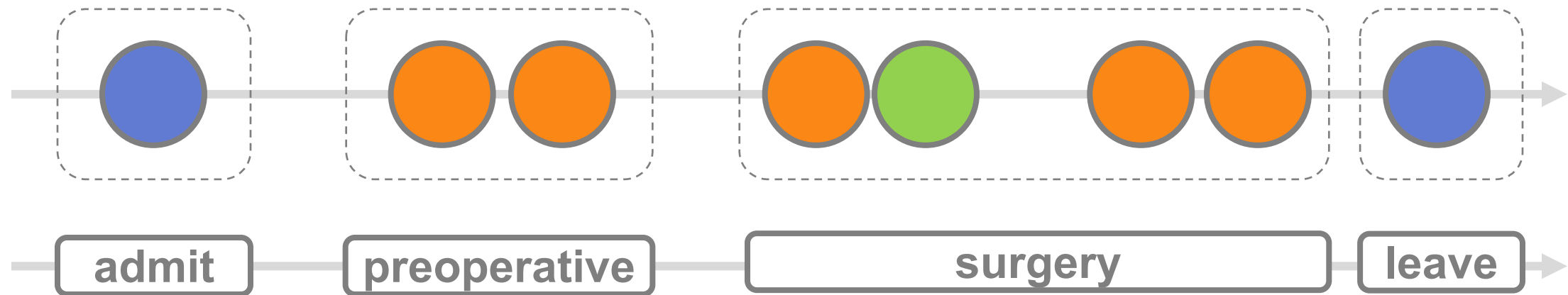
Visual Progression Analysis of Event Sequence Data

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

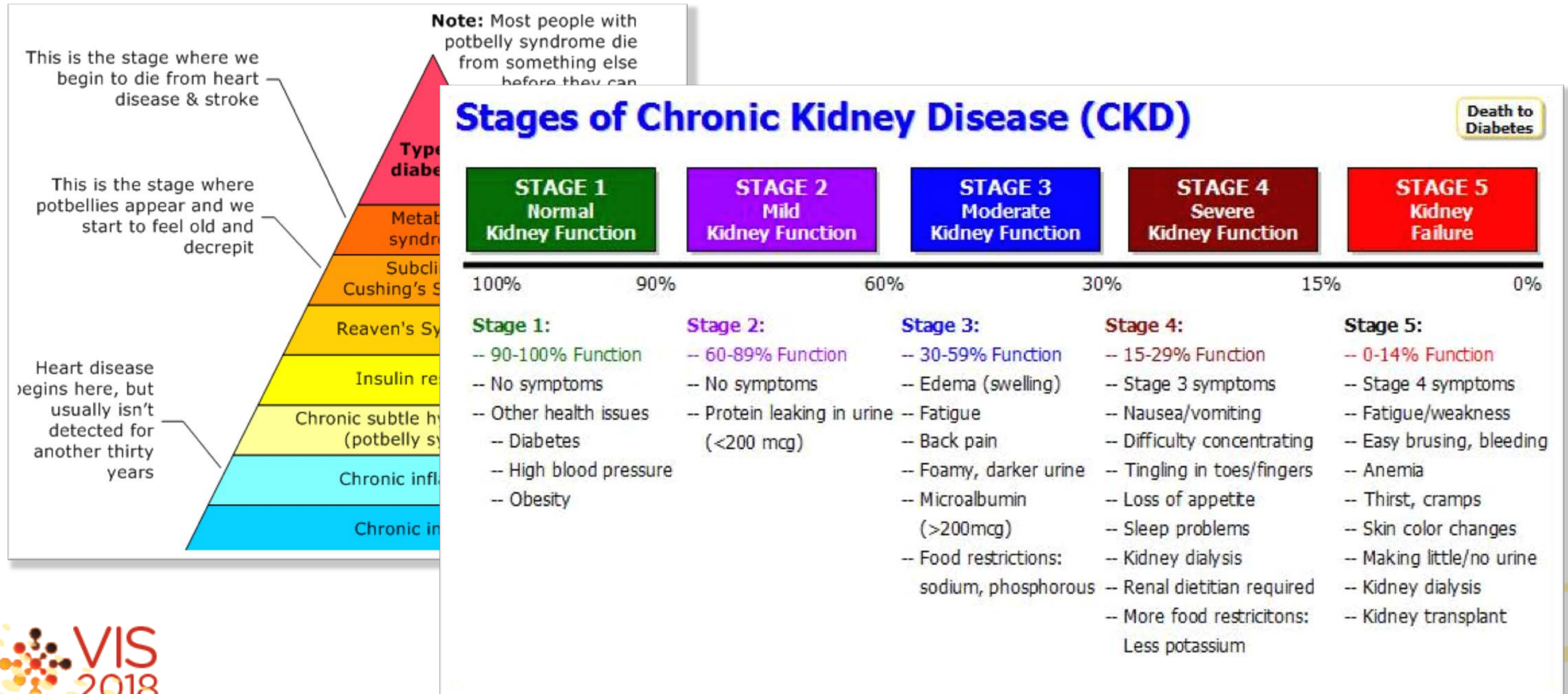
Progression Analysis on Event Sequences



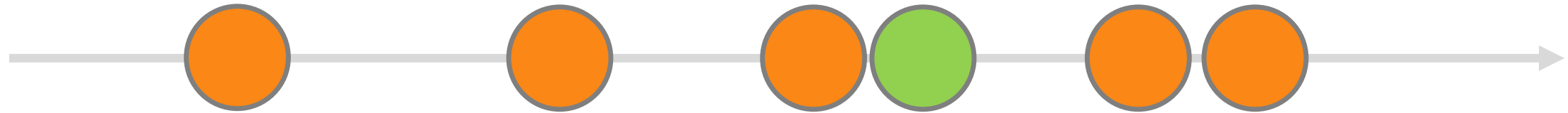
Progression of Medical Procedures



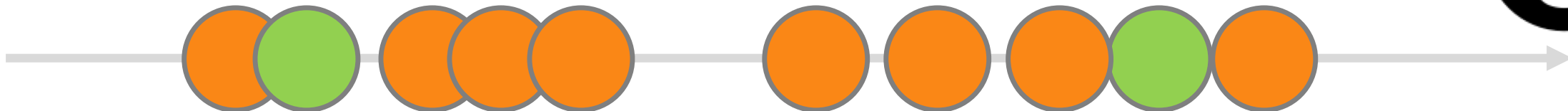
Knowing Progression Patterns is Useful



Challenges in Progression Analysis



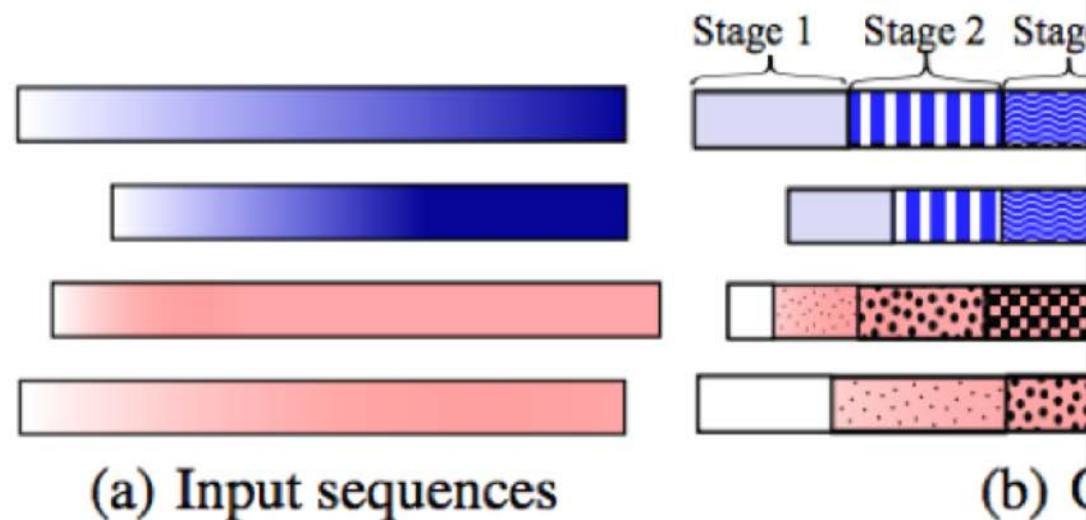
Different length / progression rate / stages



Challenges in Progression Analysis

Finding Progression Stages in Time-evolving Event Sequences

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

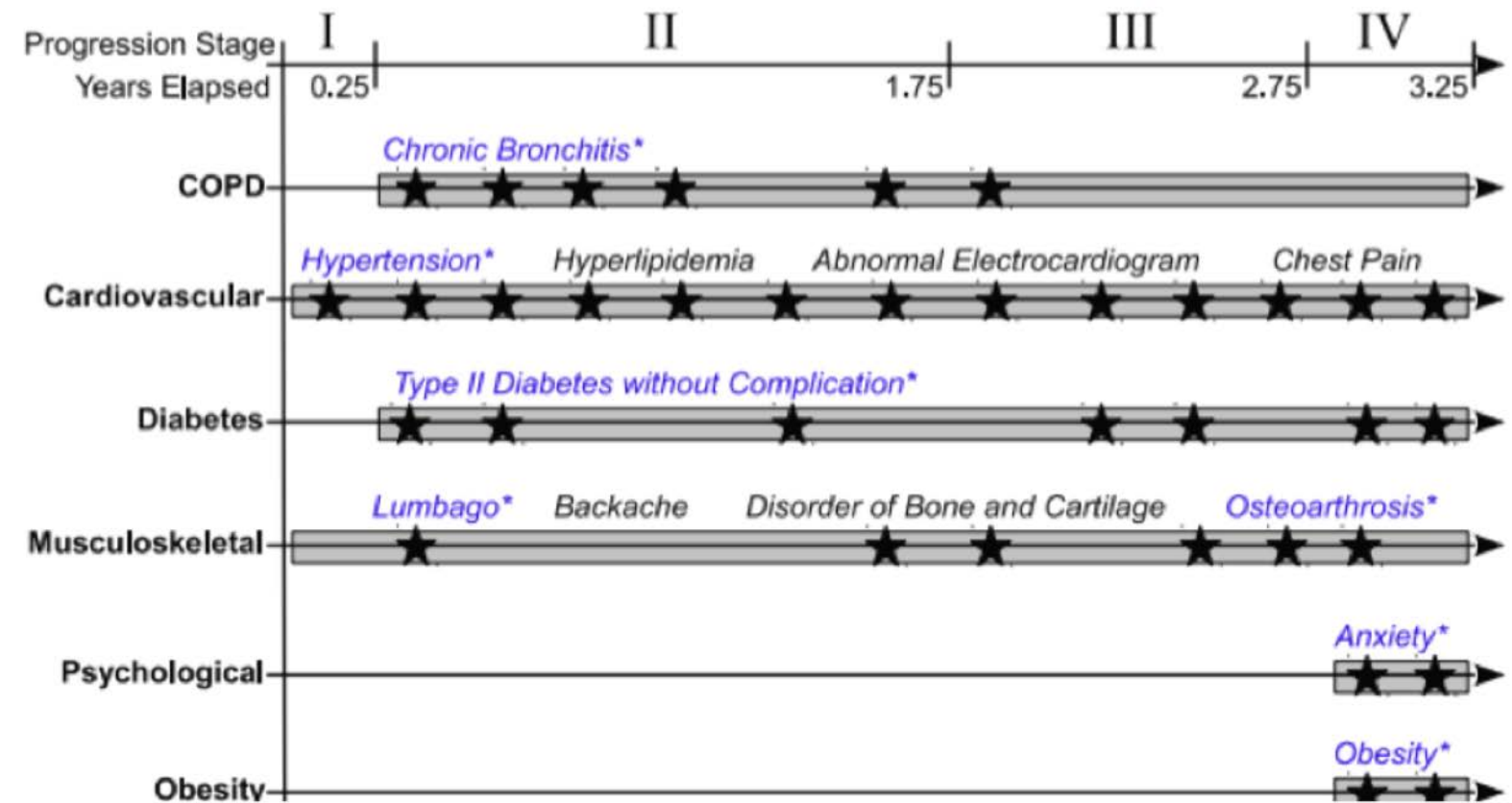


Unsupervised Learning of Disease Progression Models

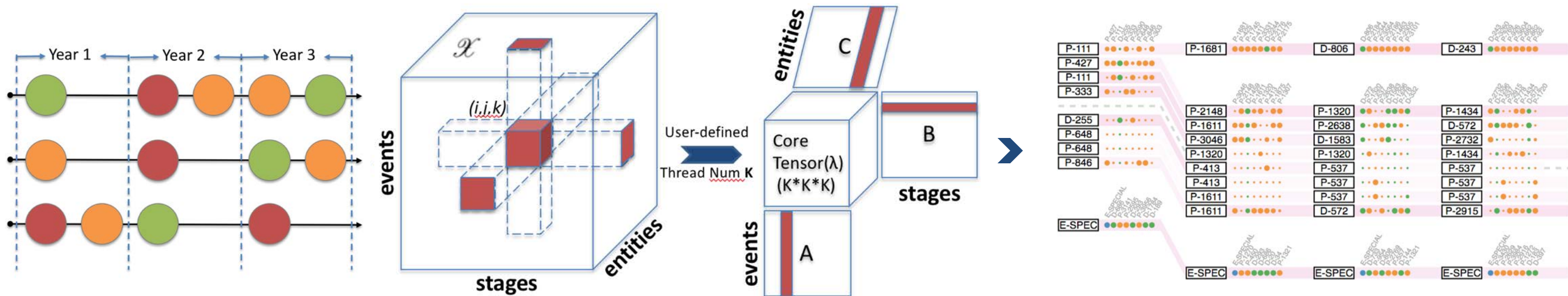
Xiang Wang
IBM Research
Yorktown Heights, NY
wangxi@us.ibm.com

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

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



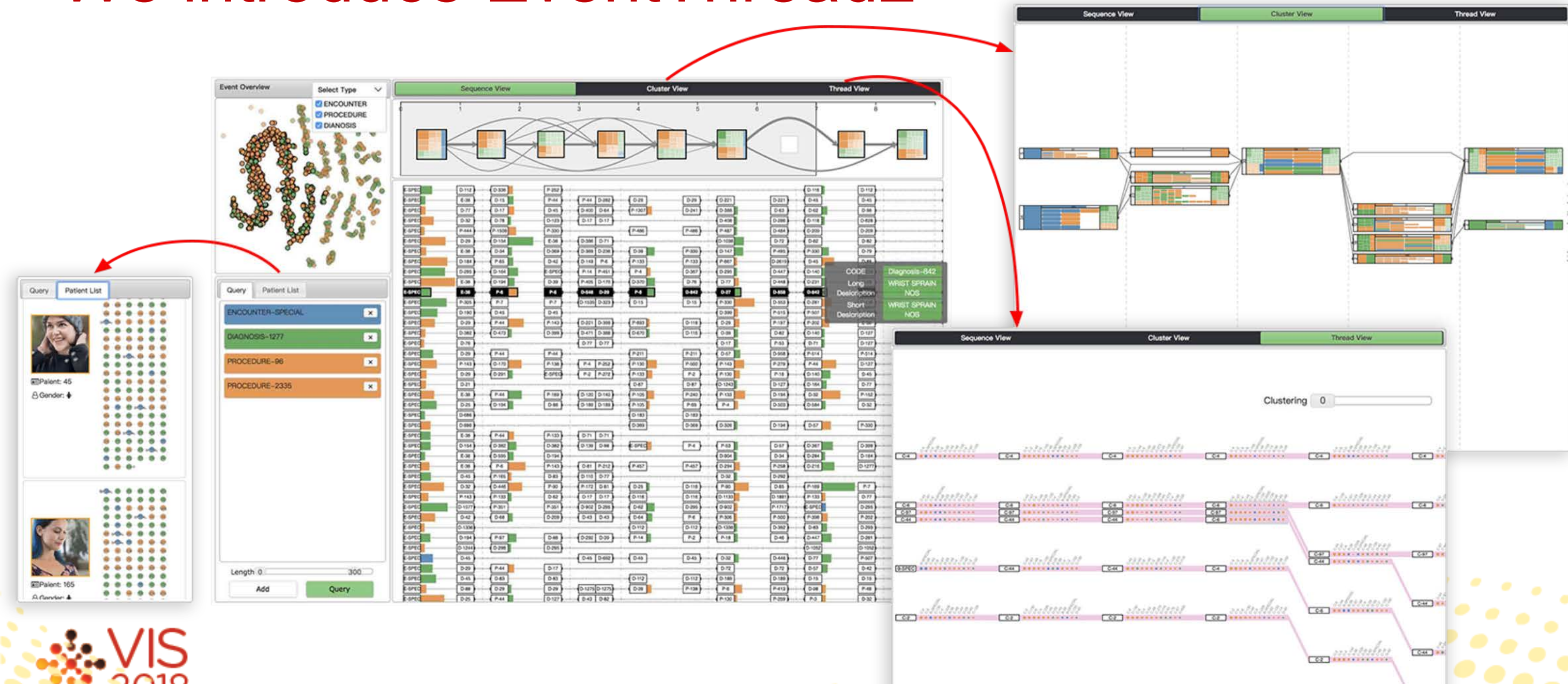
Previous Study - EventThread



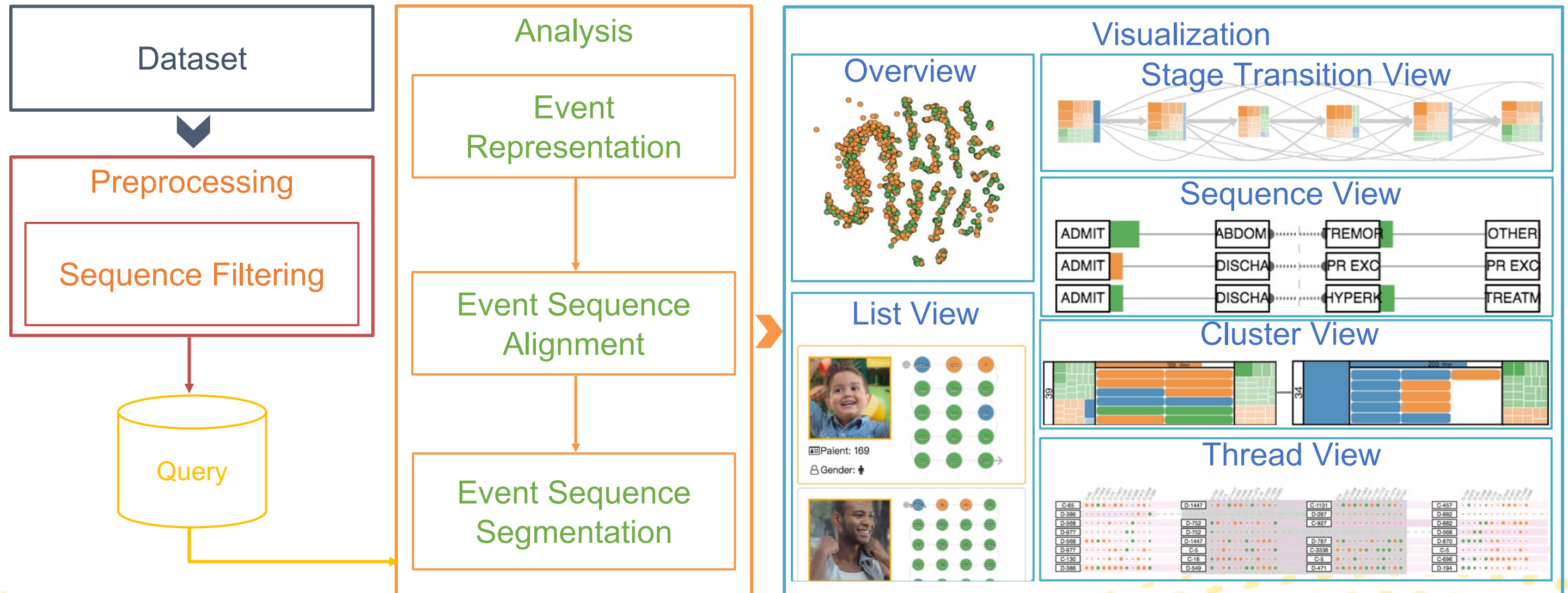
Guo, S., Xu, K., Zhao, R., Gotz, D., Zha, H., & Cao, N. (VAST'2018).

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

We introduce EventThread2

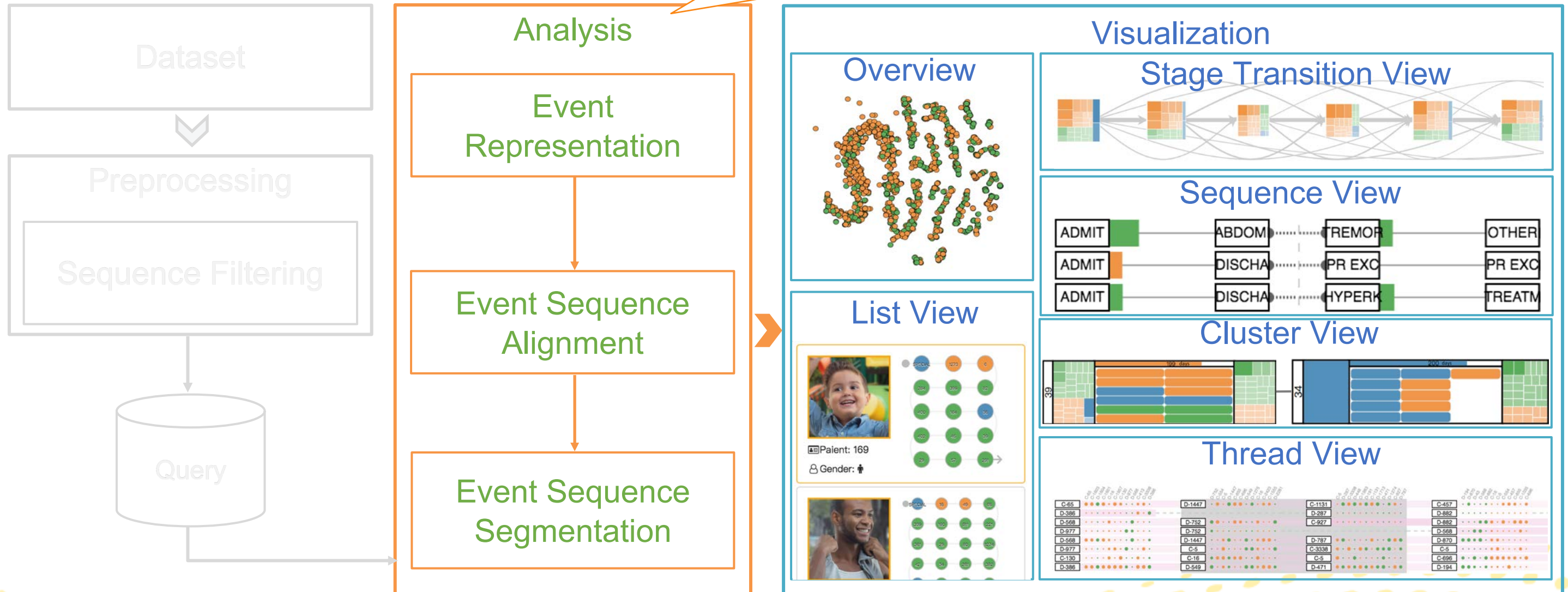


System Overview



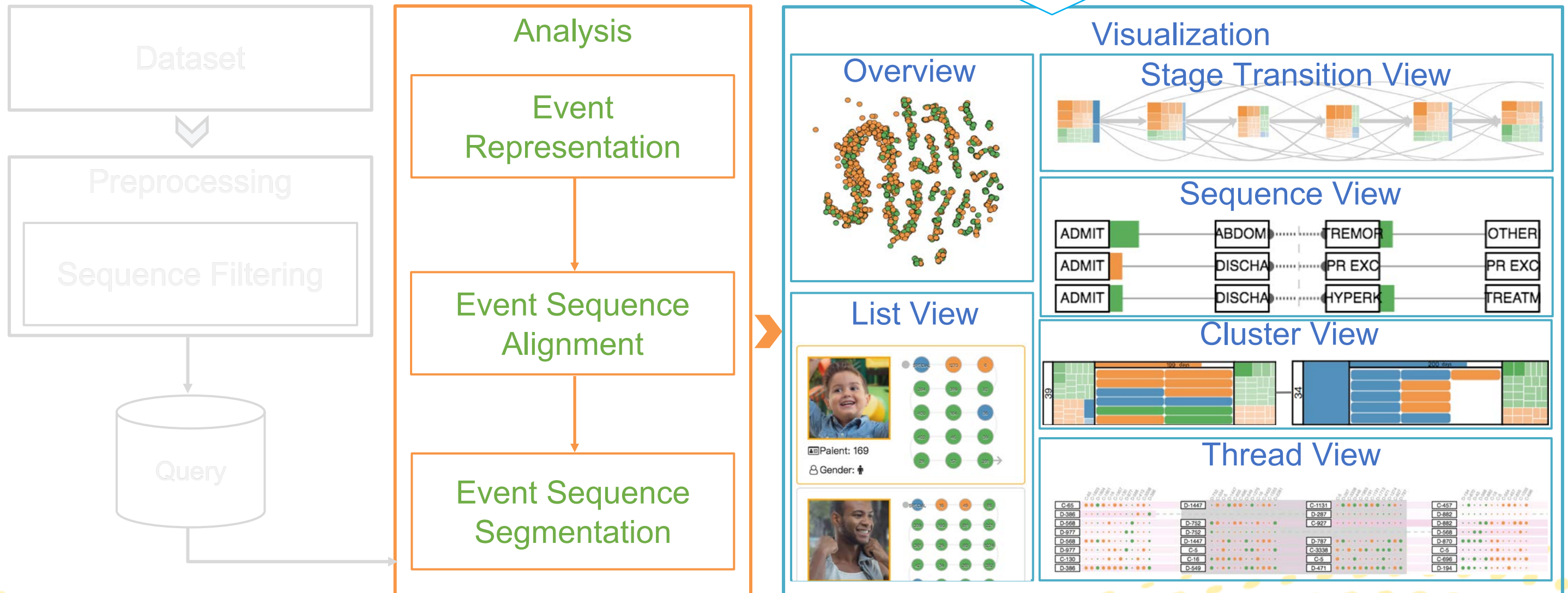
System Overview

How to identify progression stages of a sequence set?

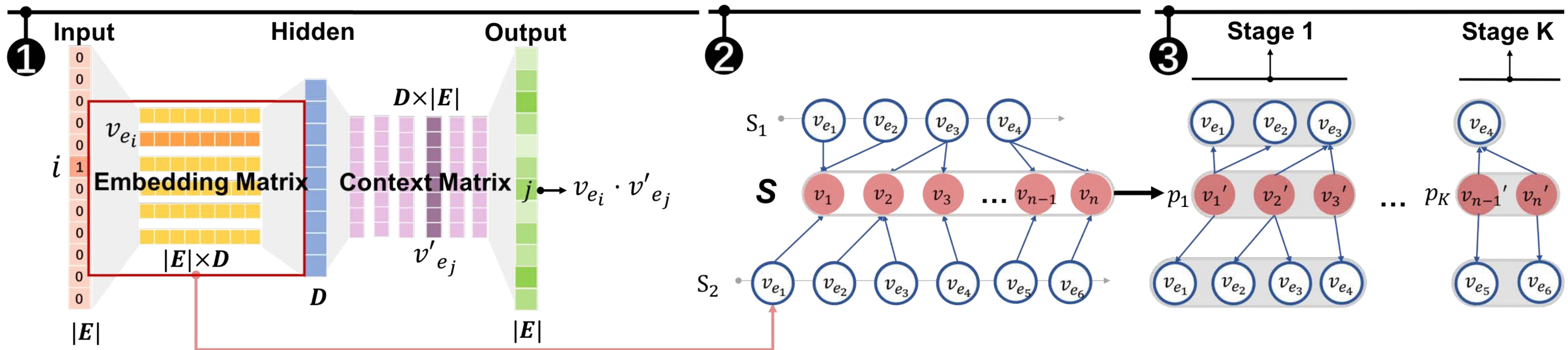


System Overview

How to visualize the result?

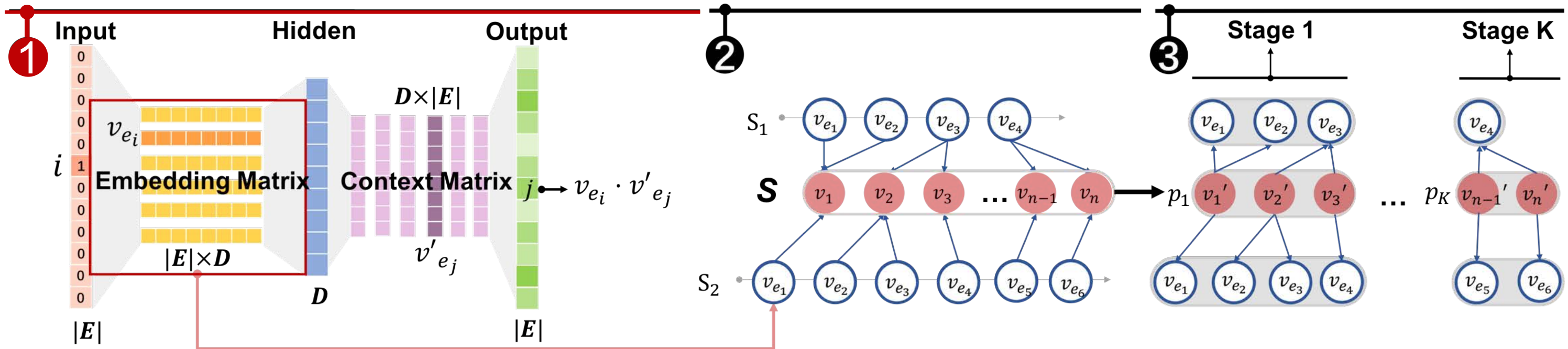


Progression Analysis

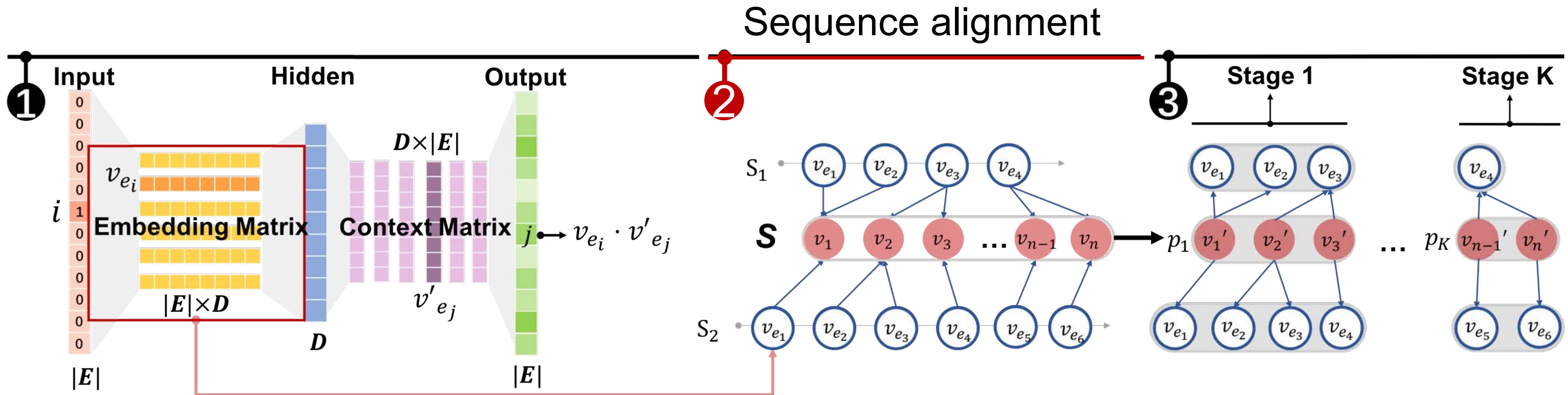


Progression Analysis

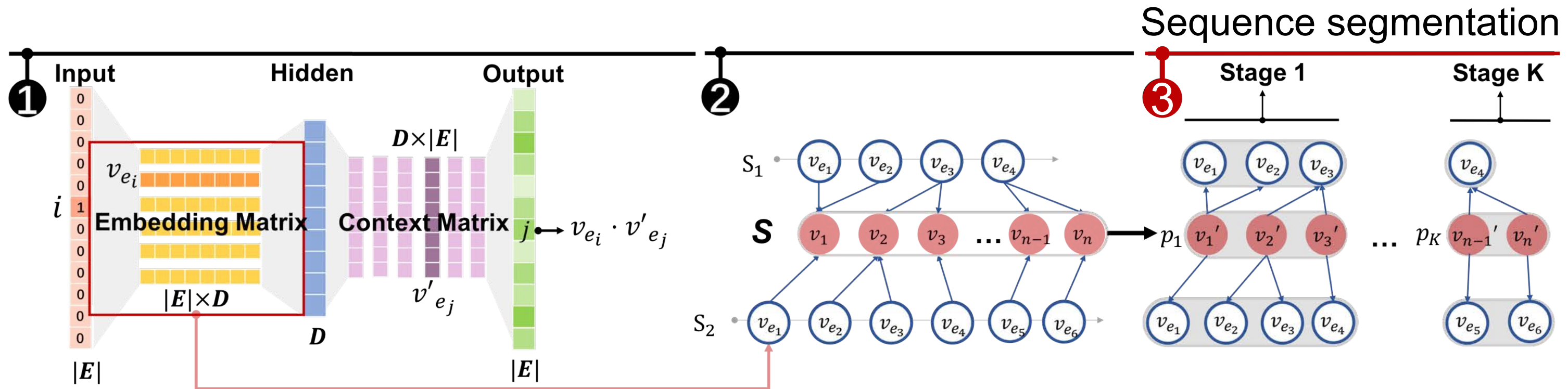
Estimation of event representation



Progression Analysis



Progression Analysis

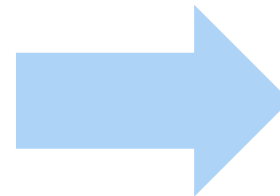


Estimate Event Representations

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

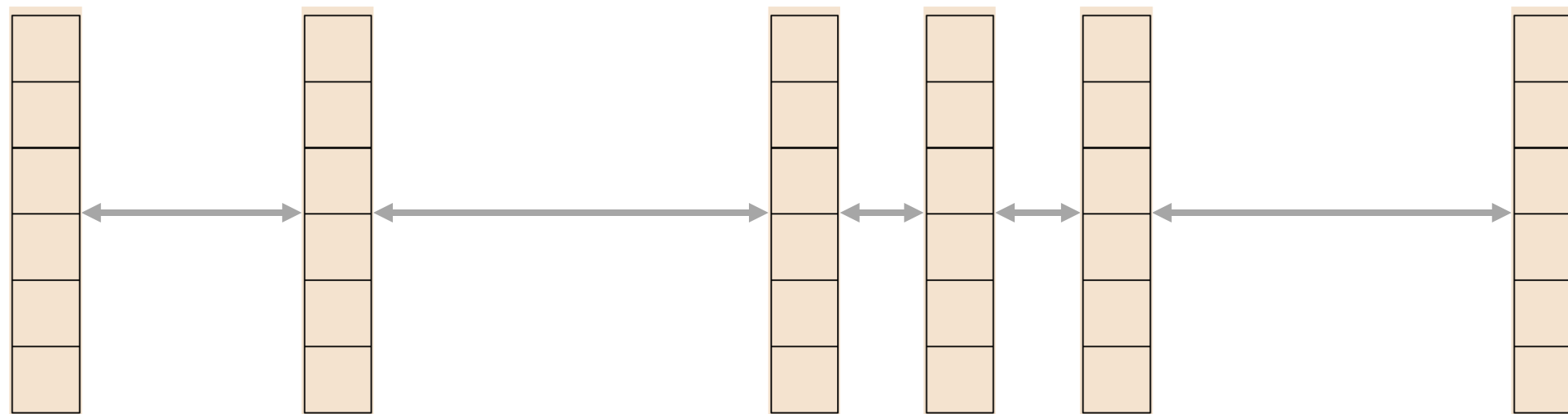
Estimate Event Representations

Which session

will you go

next

**word
embedding**

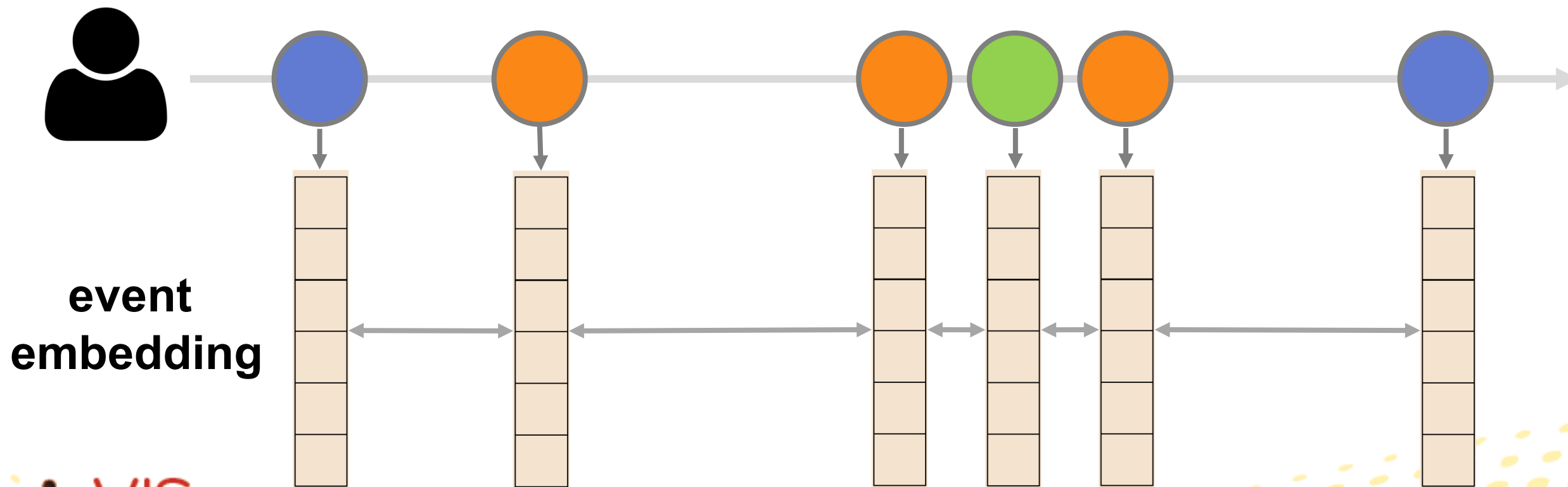


Estimate Event Representations

Which session

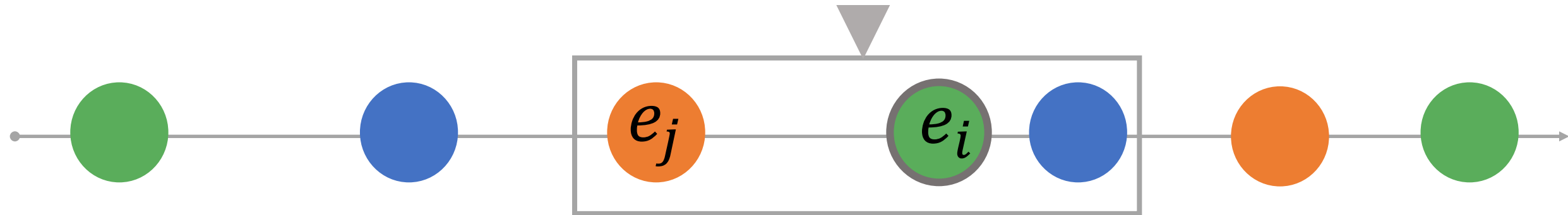
will you go

next



Estimate Event Representations

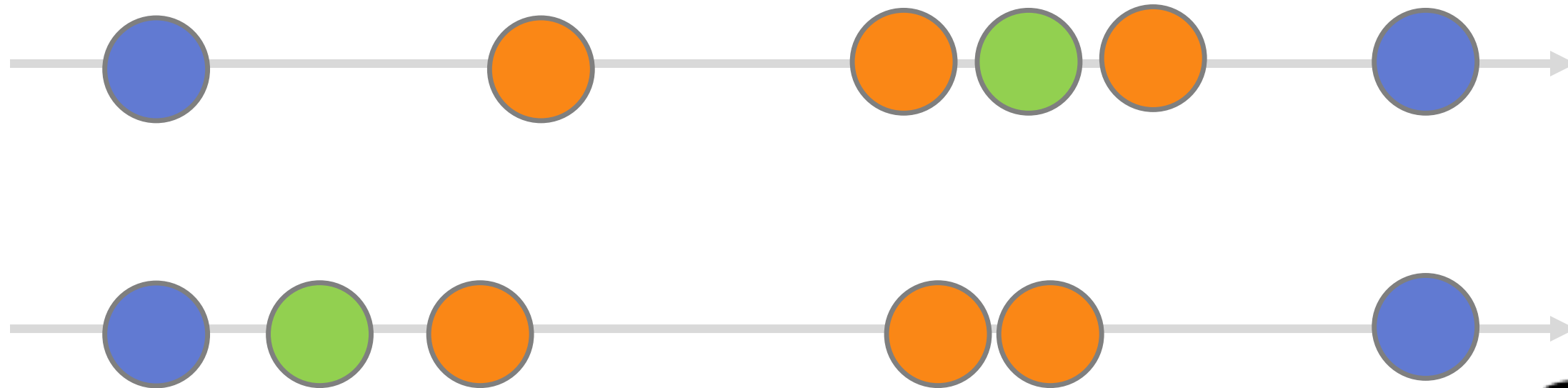
Which session will you go next



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

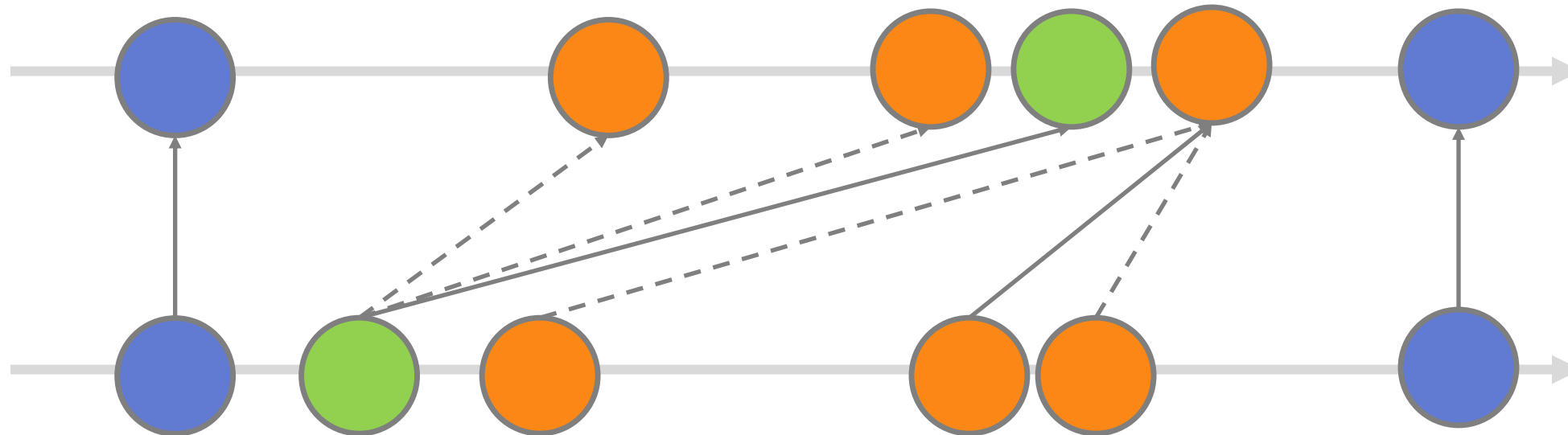


Time-based alignment



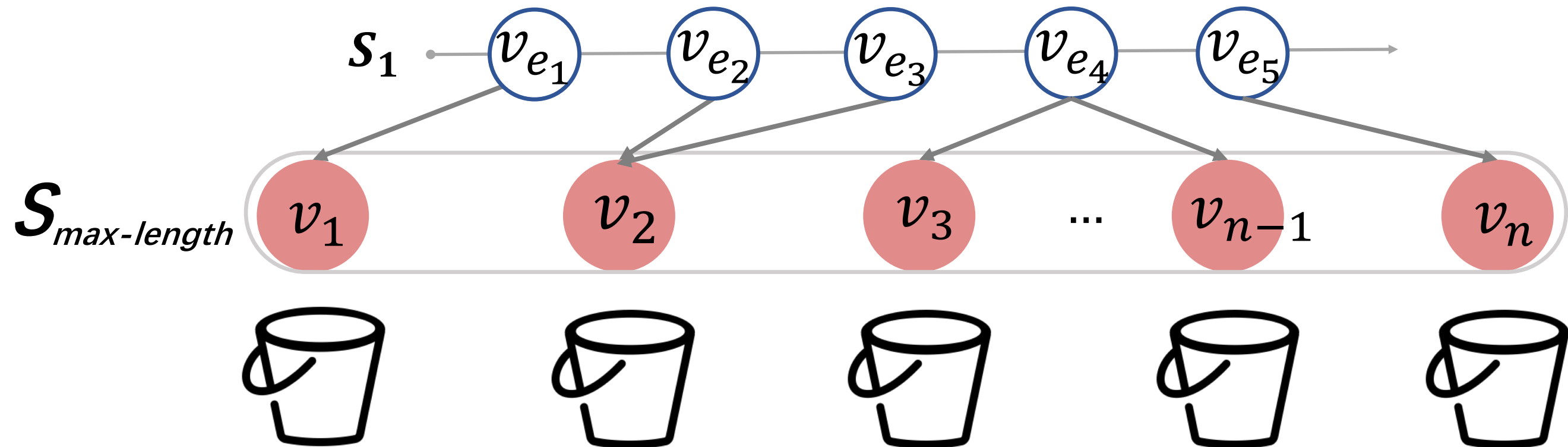
Order-based alignment

Sequence Alignment

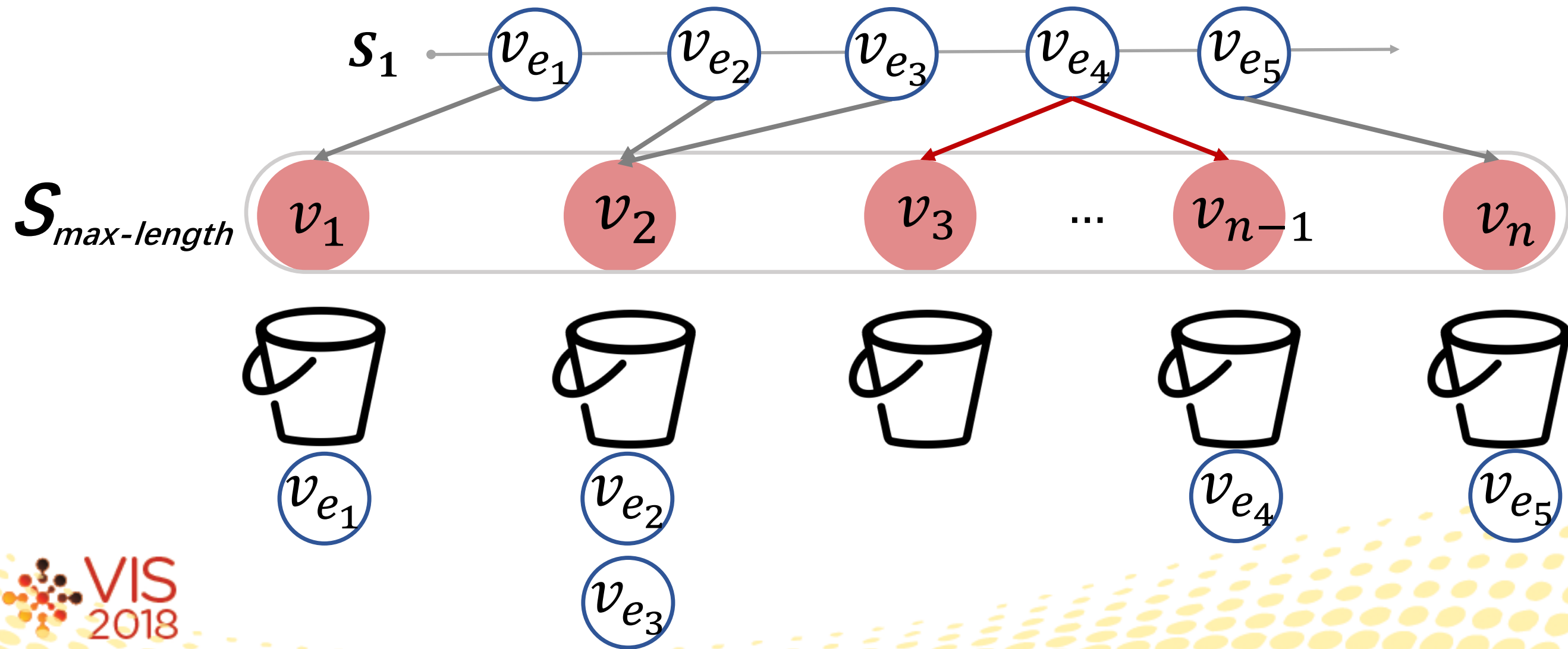


Dynamic Time Warping

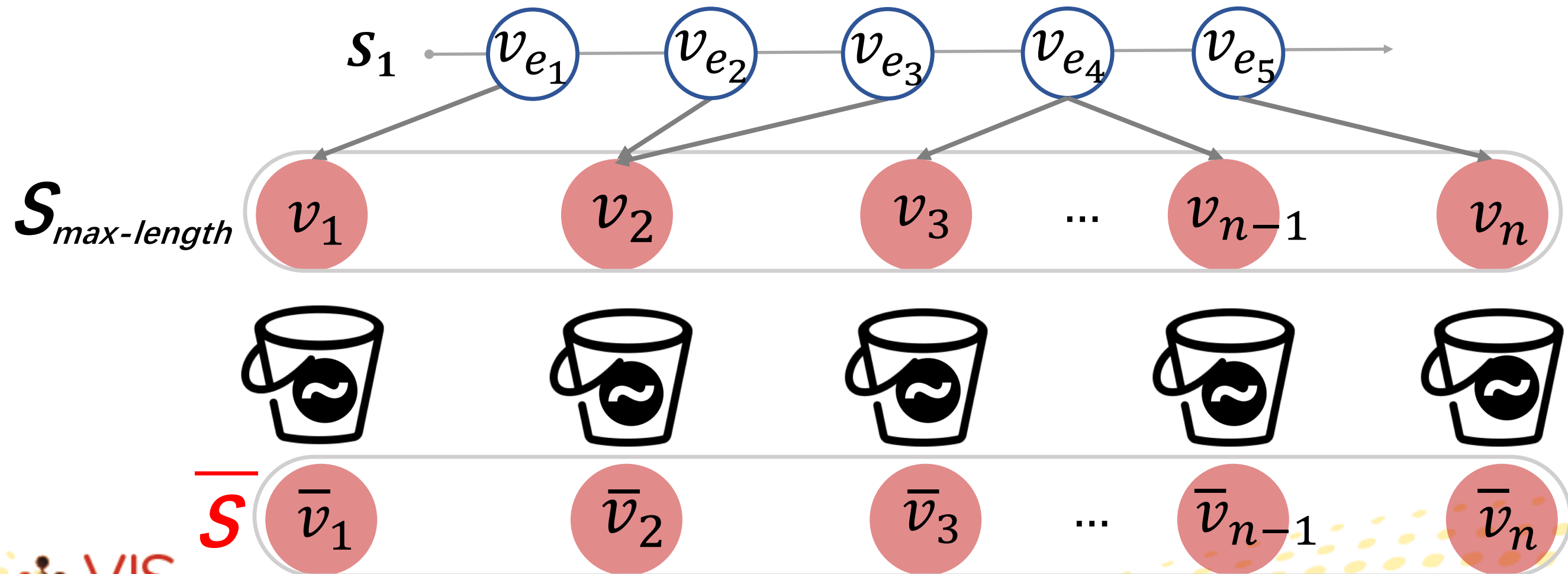
Sequence Alignment



Sequence Alignment



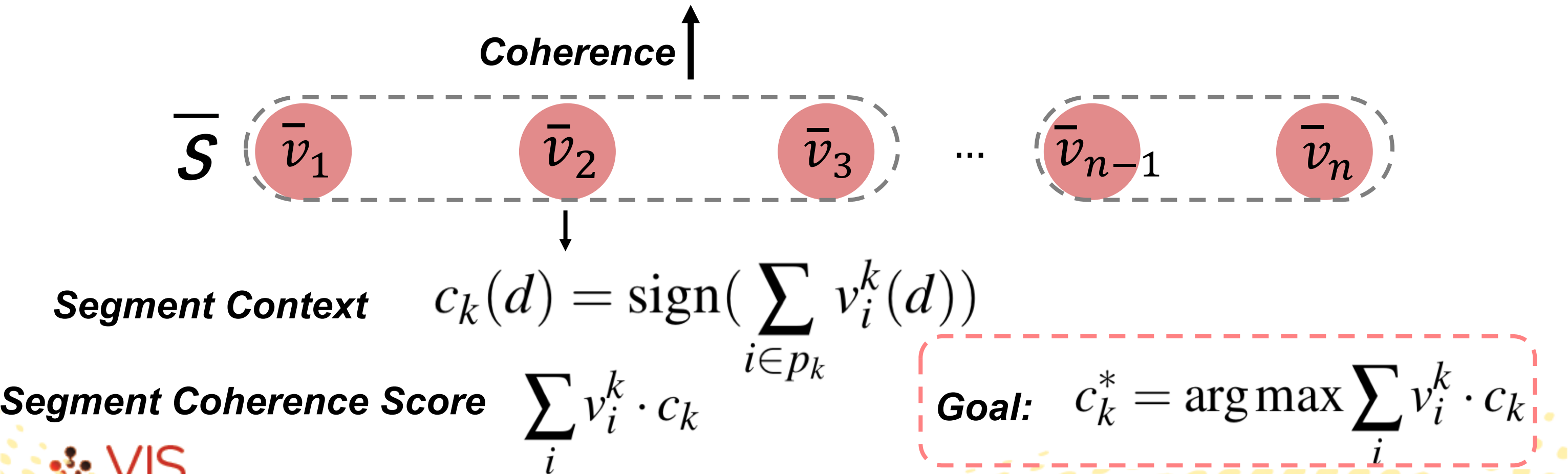
Sequence Alignment



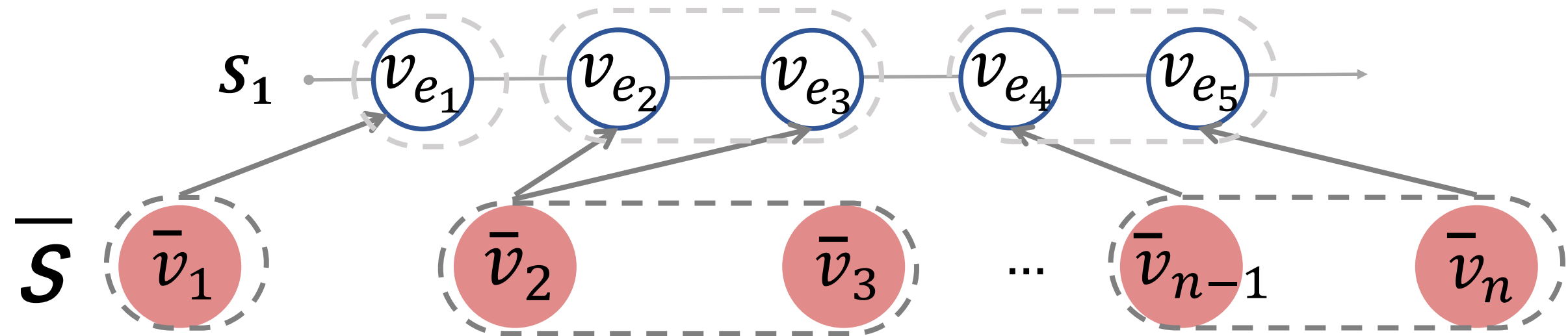
Sequence Segmentation



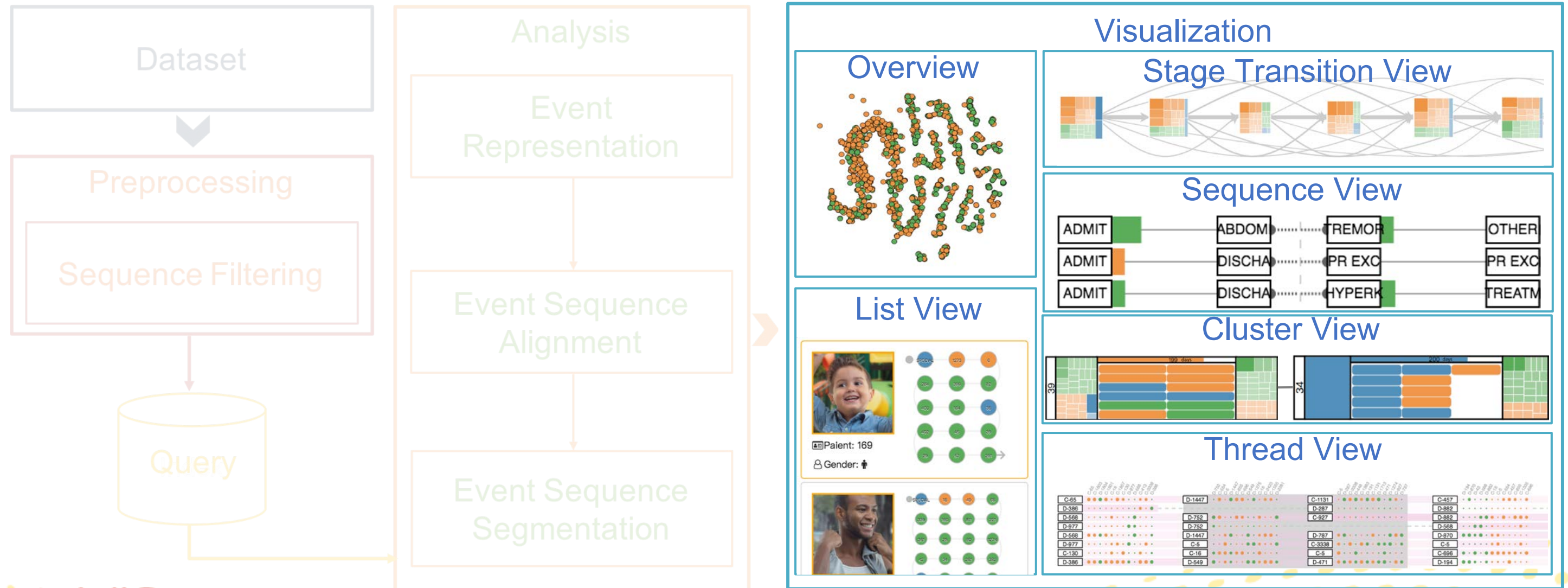
Sequence Segmentation



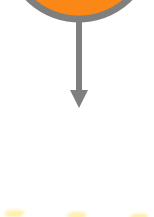
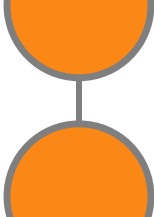
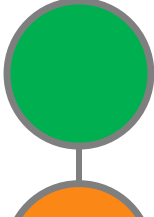
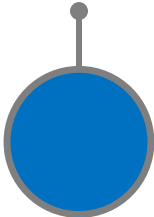
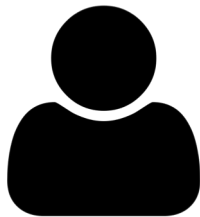
Sequence Segmentation



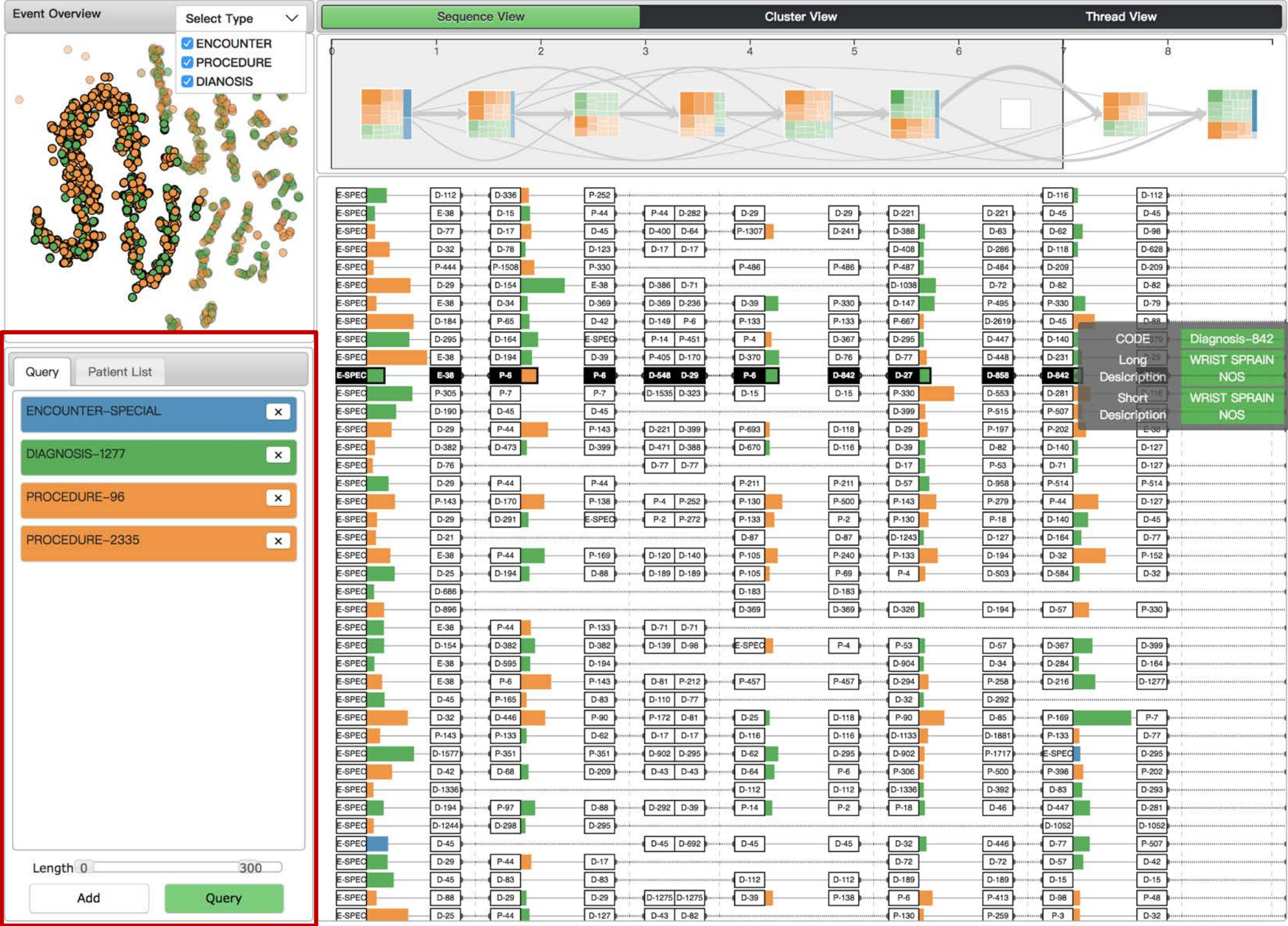
System Overview



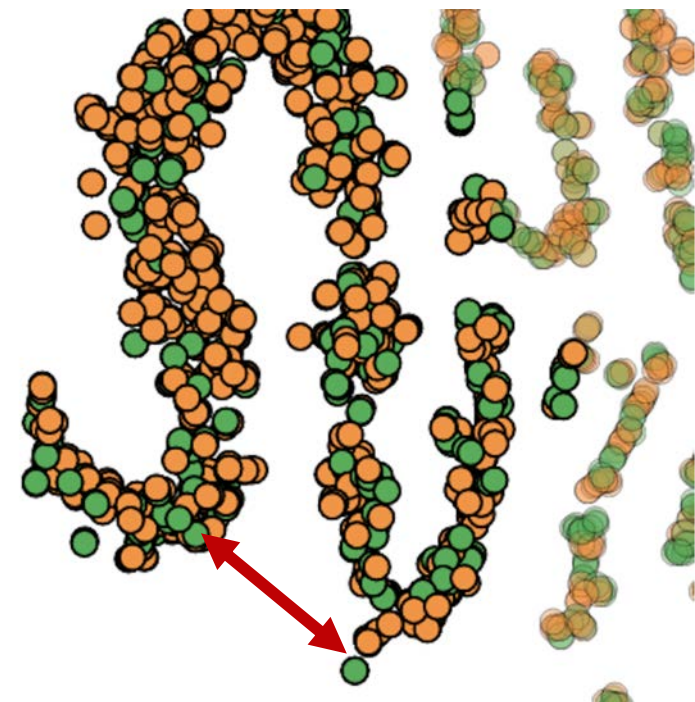
User Interface



*Milestone
Events*



User Interface



- Encounter
- Procedure
- Diagnosis

Event Overview

Select Type

ENCOUNTER

PROCEDURE

DIANOSIS

Query

Patient List

ENCOUNTER-SPECIAL

DIAGNOSIS-1277

PROCEDURE-96

PROCEDURE-2335

Length 0

300

Add

Query

Sequence View

Cluster View

Thread View

E-SPEC

D-112

D-336

P-252

P-44

D-282

D-29

D-29

D-221

D-221

D-45

D-112

E-SPEC

E-38

D-15

P-44

P-44

D-64

P-1307

D-241

D-388

D-63

D-62

D-45

E-SPEC

D-77

D-17

D-45

D-400

D-17

D-17

D-408

D-286

D-118

D-628

E-SPEC

D-32

D-78

D-123

D-17

D-17

D-486

D-486

D-484

D-209

D-209

D-82

E-SPEC

P-444

P-1508

P-330

D-386

D-71

P-486

P-486

D-1038

D-72

D-82

D-79

E-SPEC

D-29

D-154

E-38

D-369

D-236

D-39

P-330

D-147

P-495

P-330

D-82

E-SPEC

E-38

D-34

D-42

D-149

P-6

P-133

P-667

D-2619

P-330

D-82

E-SPEC

D-184

P-65

E-SPEC

P-14

P-451

P-4

D-367

D-295

D-447

D-140

D-82

E-SPEC

D-295

D-164

D-39

P-405

D-170

D-370

D-76

D-77

D-448

D-231

D-82

E-SPEC

E-38

D-194

D-39

P-405

D-170

D-370

D-76

D-77

D-448

D-231

D-82

E-SPEC

E-38

D-194

D-39

P-405

D-170

D-370

D-76

D-77

D-448

D-231

D-82

E-SPEC

P-305

P-7

P-7

D-1535

D-323

D-15

D-15

P-330

D-553

P-507

D-281

E-SPEC

D-190

D-45

D-45

D-45

D-399

P-693

D-118

D-29

P-197

P-202

D-127

E-SPEC

D-29

P-44

P-143

D-221

D-399

P-693

D-118

D-29

P-197

P-202

D-127

E-SPEC

D-382

D-473

D-399

D-471

D-388

D-670

D-116

D-39

D-82

D-140

D-127

E-SPEC

D-76

D-77

D-77

D-77

D-77

D-77

D-116

D-39

D-82

D-140

D-127

E-SPEC

D-29

P-44

P-44

P-44

P-211

P-211

D-57

D-958

P-514

P-514

D-127

E-SPEC

P-143

D-170

P-138

P-4

P-252

P-130

P-500

P-143

P-279

P-44

D-127

E-SPEC

D-29

D-291

E-SPEC

P-2

P-272

P-133

P-2

P-130

P-18

D-140

D-140

D-45

E-SPEC

D-21

P-44

P-169

D-120

D-140

P-105

P-240

P-133

D-127

D-164

D-77

E-SPEC

E-38

P-44

P-169

D-120

D-140

P-105

P-240

P-133

D-127

D-164

D-77

E-SPEC

D-25

D-194

D-88

D-189

D-189

P-105

P-69

P-4

D-503

D-584

D-32

E-SPEC

D-686

D-88

D-189

D-189

P-105

P-69

P-4

D-503

D-584

D-32

E-SPEC

D-896

D-88

D-189

D-189

P-105

P-69

P-4

D-503

D-584

D-32

E-SPEC

E-38

P-44

P-133

D-71

D-71

E-SPEC

P-4

P-53

D-57

D-367

D-399

E-SPEC

D-154

D-382

D-382

D-139

D-98

E-SPEC

P-4

P-53

D-57

D-367

D-399

E-SPEC

E-38

D-595

D-194

D-194

D-81

P-212

P-457

P-457

D-294

P-258

D-216

D-1277

E-SPEC

E-38

P-6

P-143

D-81

P-212

P-457

P-457

D-294

P-258

D-216

D-1277

E-SPEC

D-45

P-165

D-83

D-110

D-77

D-118

D-118

D-32

D-292

D-85

P-169

P-7

E-SPEC

D-32

D-446

P-90

P-172

D-81

D-25

D-118

D-118

D-32

D-292

D-85

P-169

P-7

E-SPEC

P-143

P-133

D-62

D-17

D-17

D-116

D-116

D-1133

D-1881

P-133

D-77

E-SPEC

D-1577

P-351

D-351

D-902

D-295

D-62

D-295

D-902

P-1717

E-SPEC

D-295

E-SPEC

D-42

D-68

D-209

D-43

D-43

D-64

P-6

P-306

P-500

P-398

P-202

E-SPEC

D-1336

D-112

D-112

D-1336

D-392

D-83

D-392

D-293

E-SPEC

D-194

P-97

D-88

D-292

D-39

P-14

P-2

P-18

D-46

D-447

D-281

E-SPEC

D-1244

D-298

D-295

D-295

D-45

D-692

D-45

D-45

D-32

D-446

D-77

P-507

E-SPEC

D-45

P-44

D-17

D-17

D-112

D-112

D-72

D-72

D-57

D-42

E-SPEC

D-45

D-83

D-83

D-112

D-112

D-189

D-189

D-15

D-15

E-SPEC

D-88

D-29

D-29

D-1275

D-1275

D-39

P-138

P-6

P-413

D-98

P-48

E-SPEC

D-25

P-44

D-127

D-43

D-82

P-130

P-259

P-3

D-32

CODE-842

Long Description

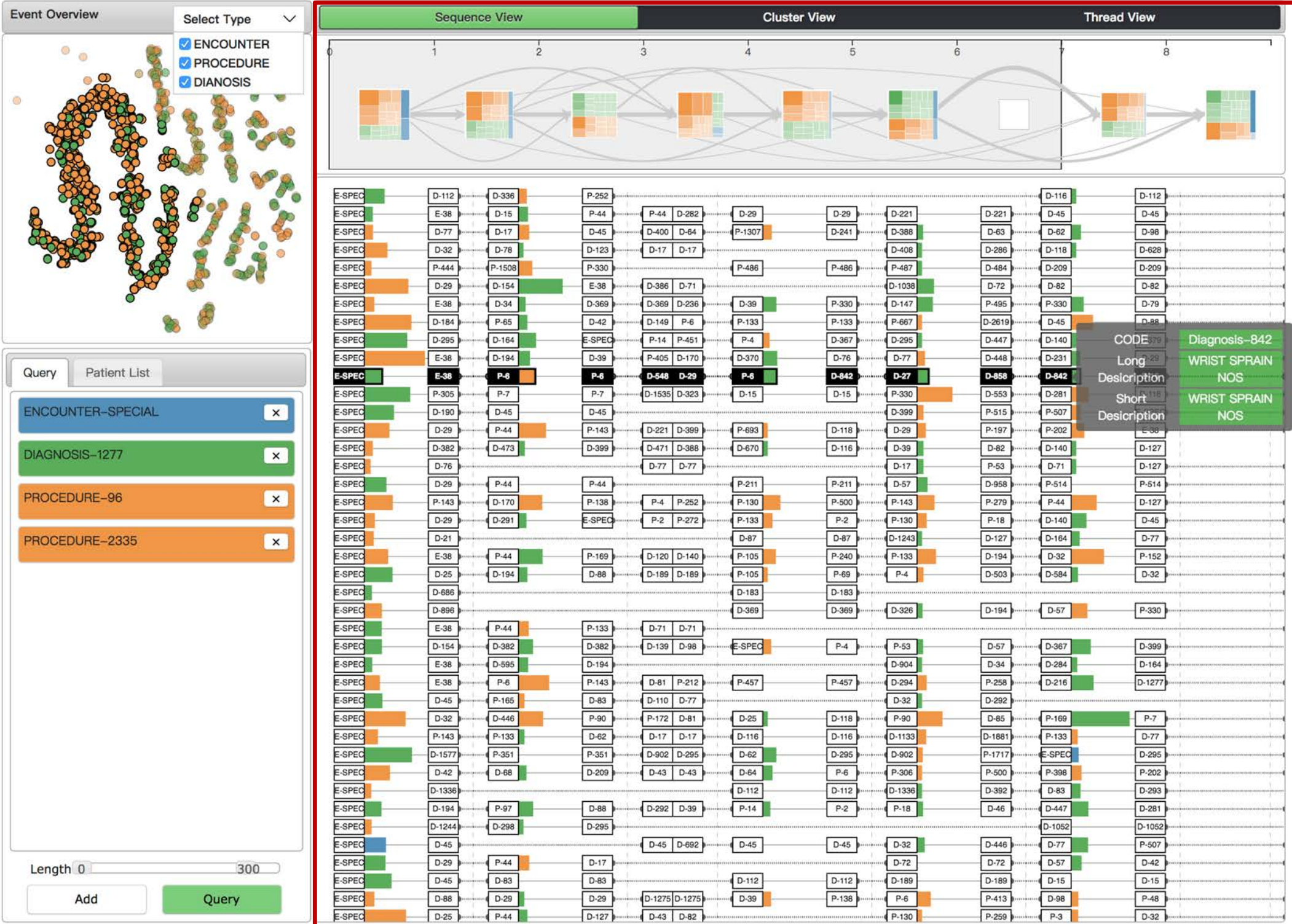
Short Description

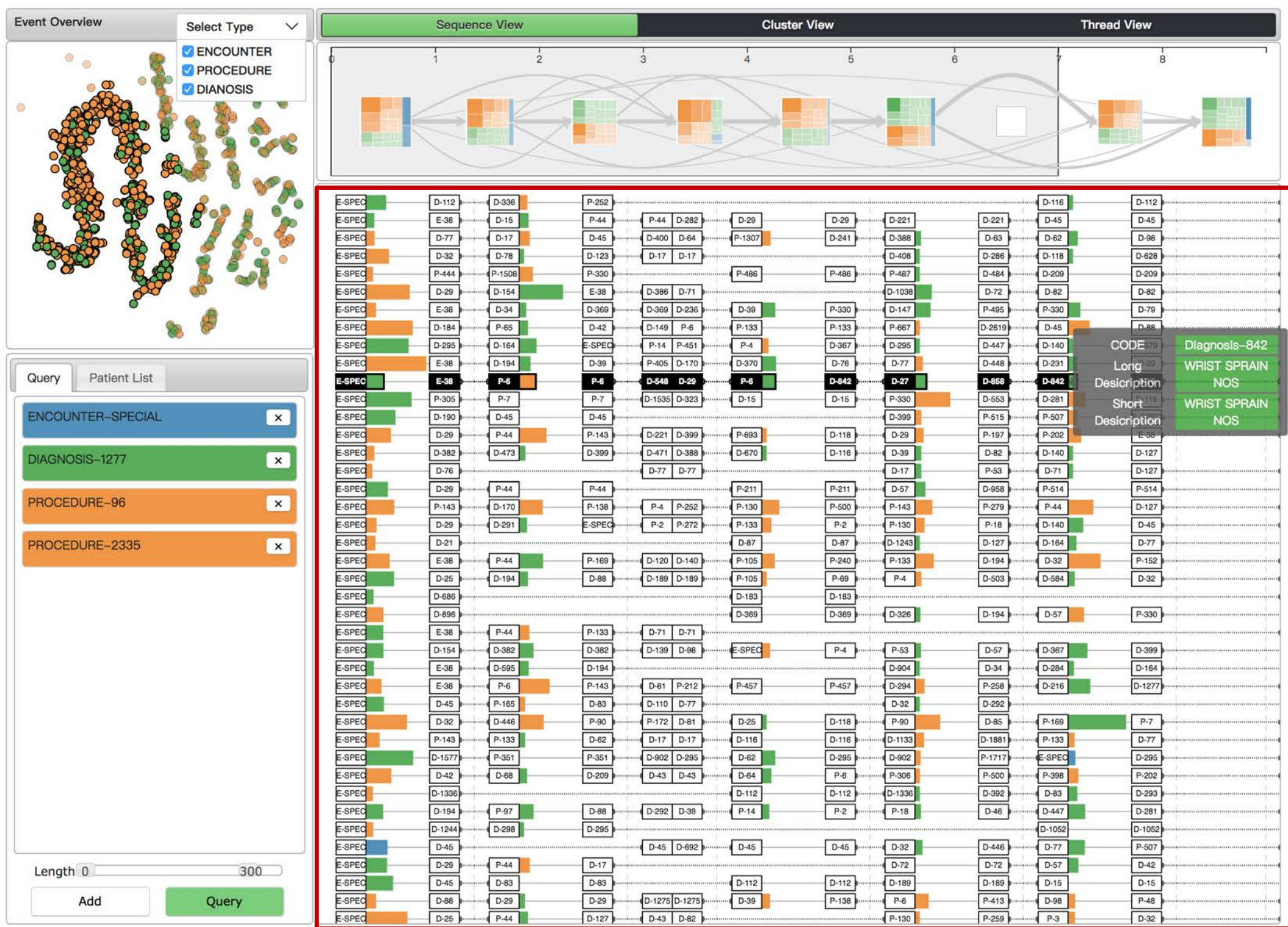
Diagnosis-842

WRIST SPRAIN NOS

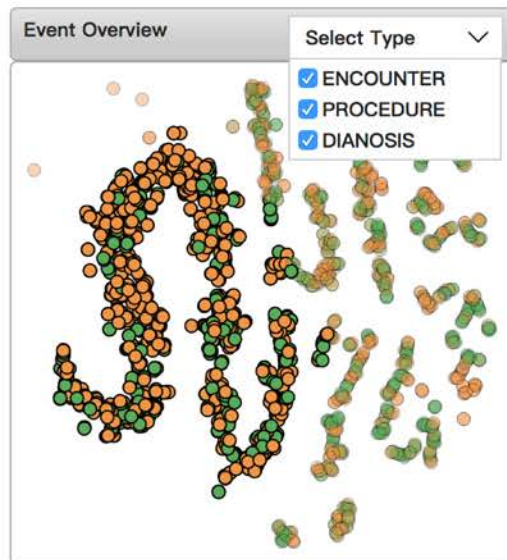
WRIST SPRAIN NOS

User Interface





Sequence view



Query Patient List

ENCOUNTER-SPECIAL x

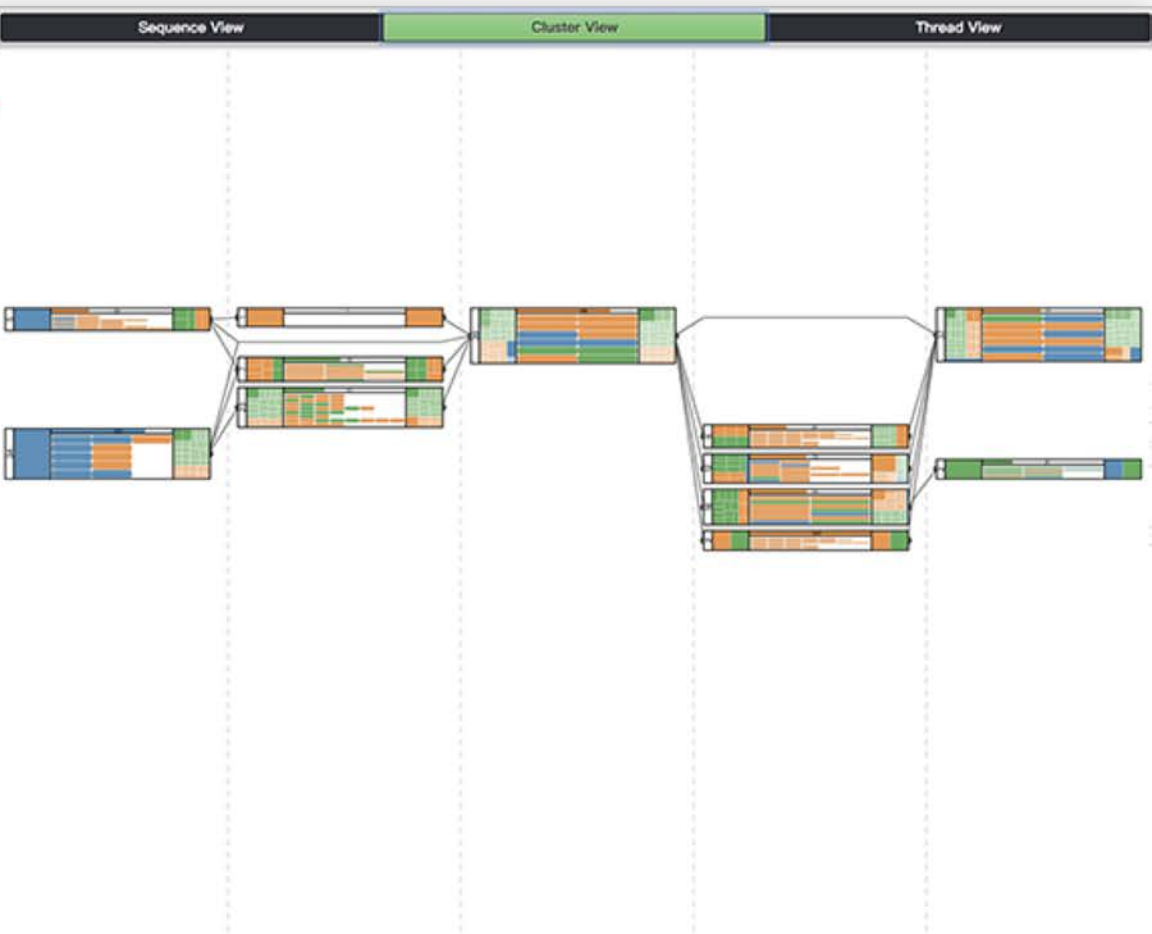
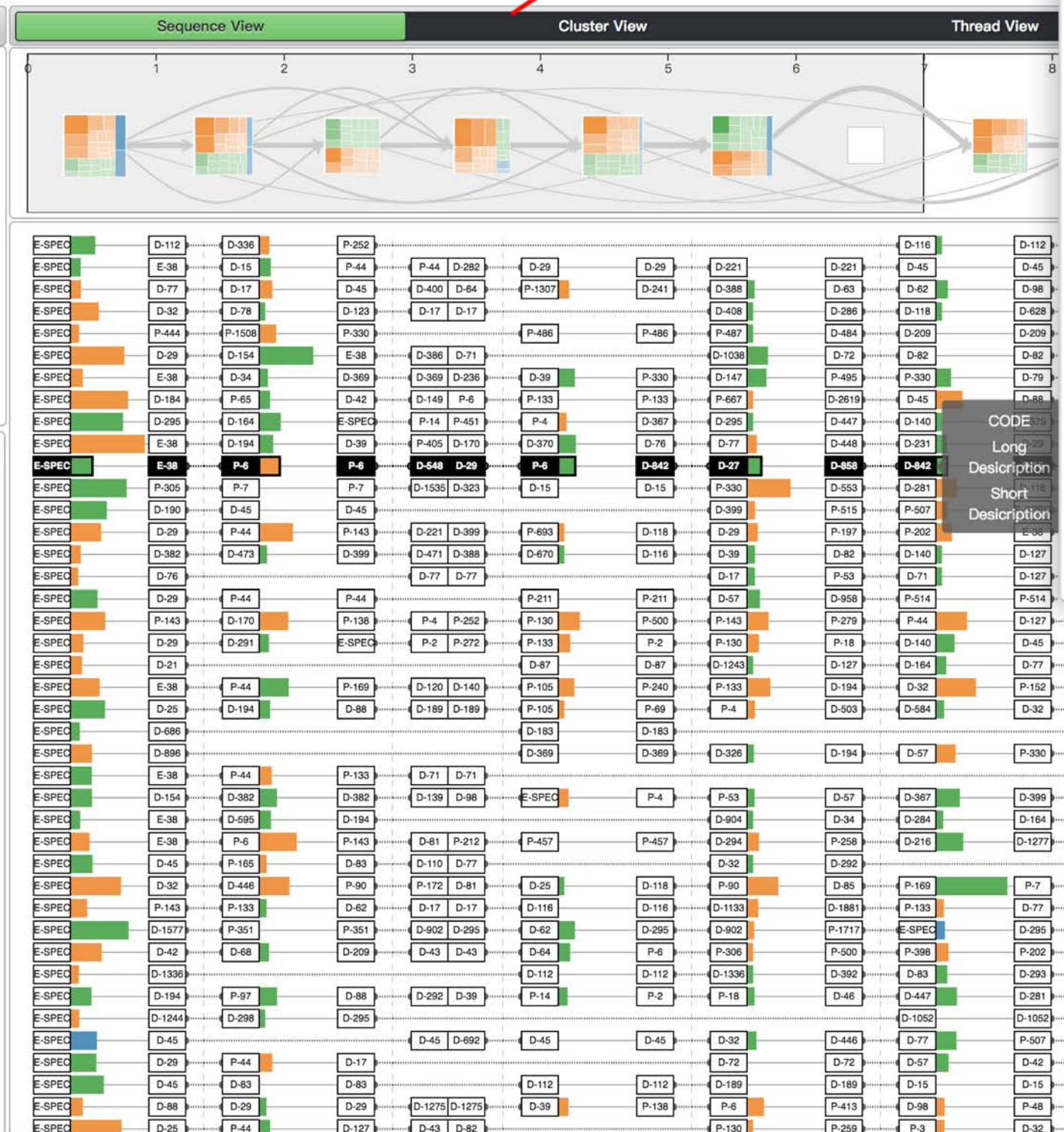
DIAGNOSIS-1277 x

PROCEDURE-96 x

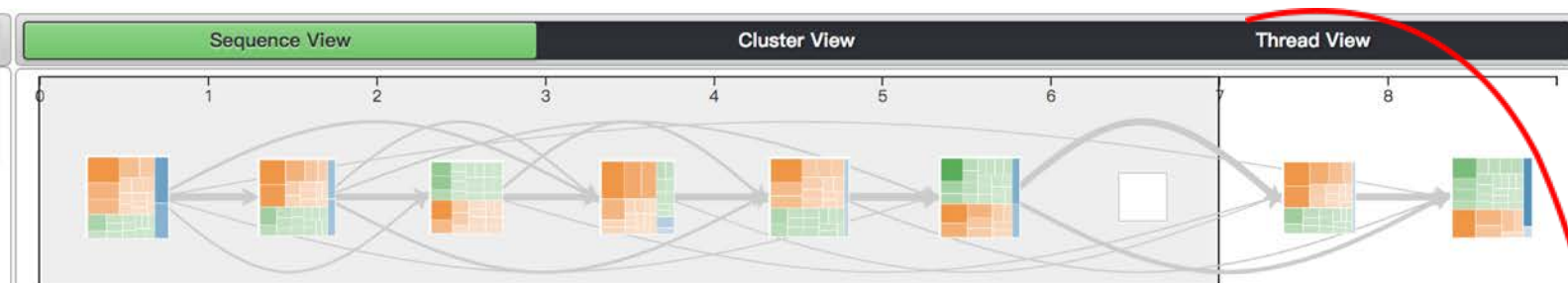
PROCEDURE-2335 x

Length 0 300

Add Query



Cluster view



Thread view

Query

Patient List

ENCOUNTER-SPECIAL

x

DIAGNOSIS-1277

x

PROCEDURE-96

x

PROCEDURE-2335

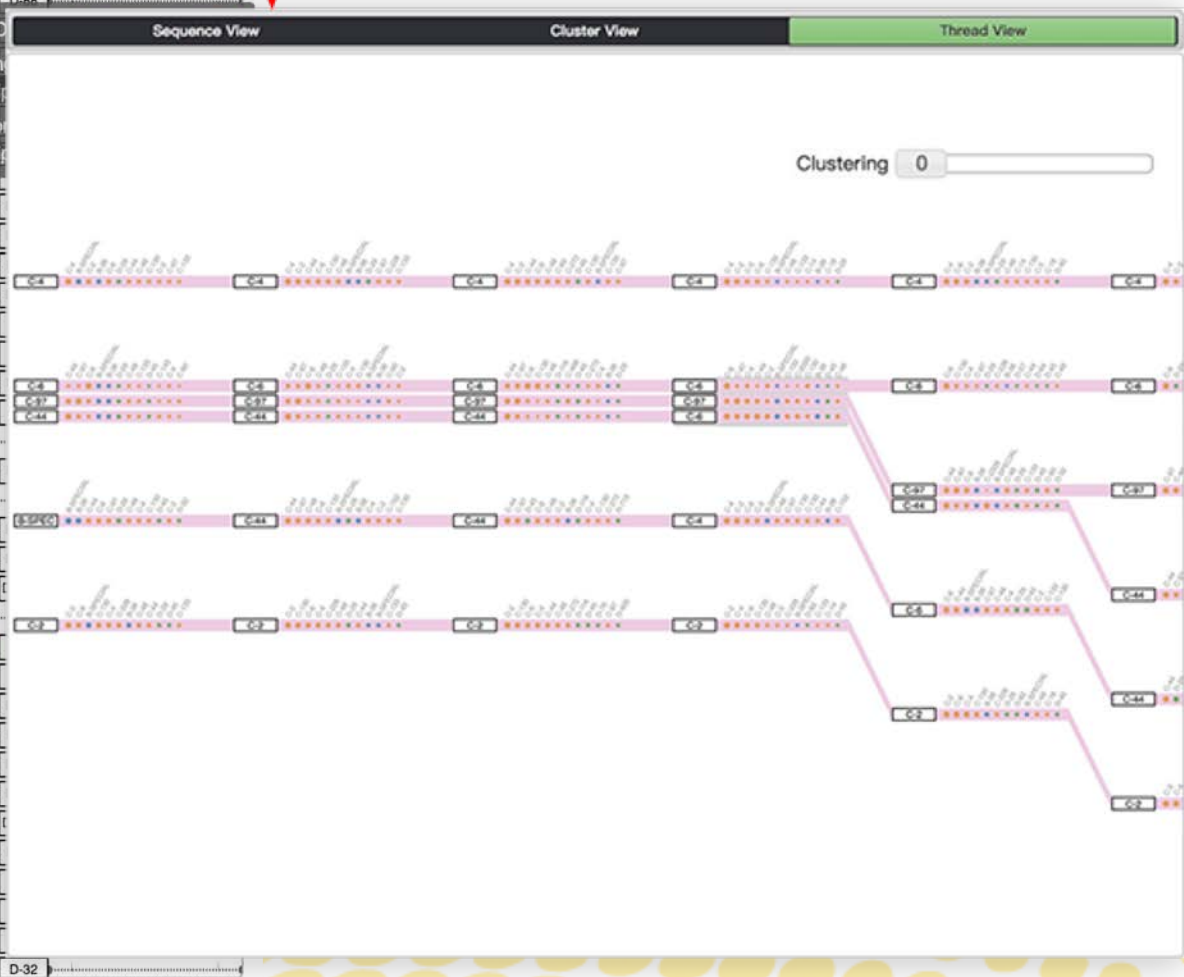
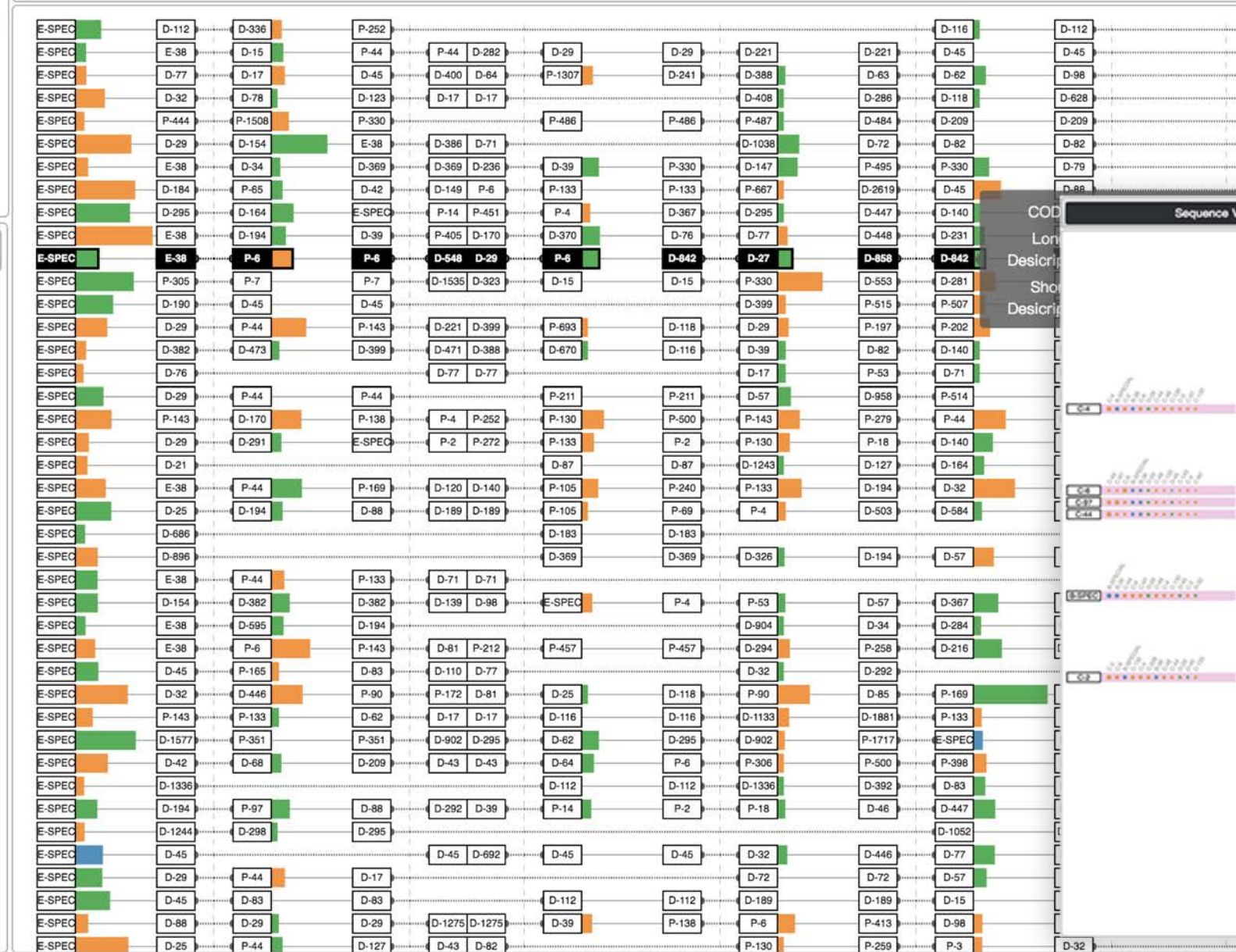
x

Length 0

300

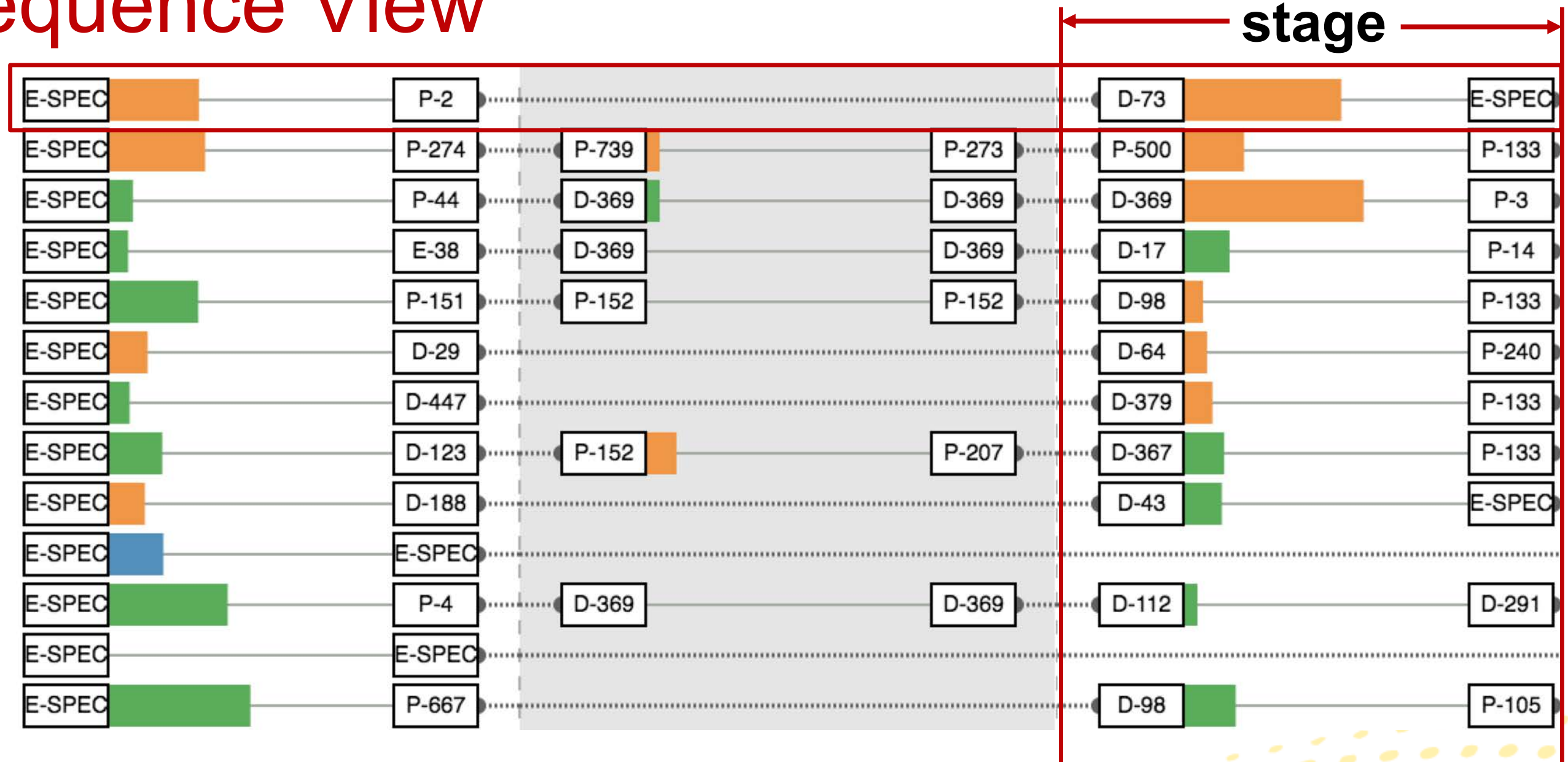
Add

Query

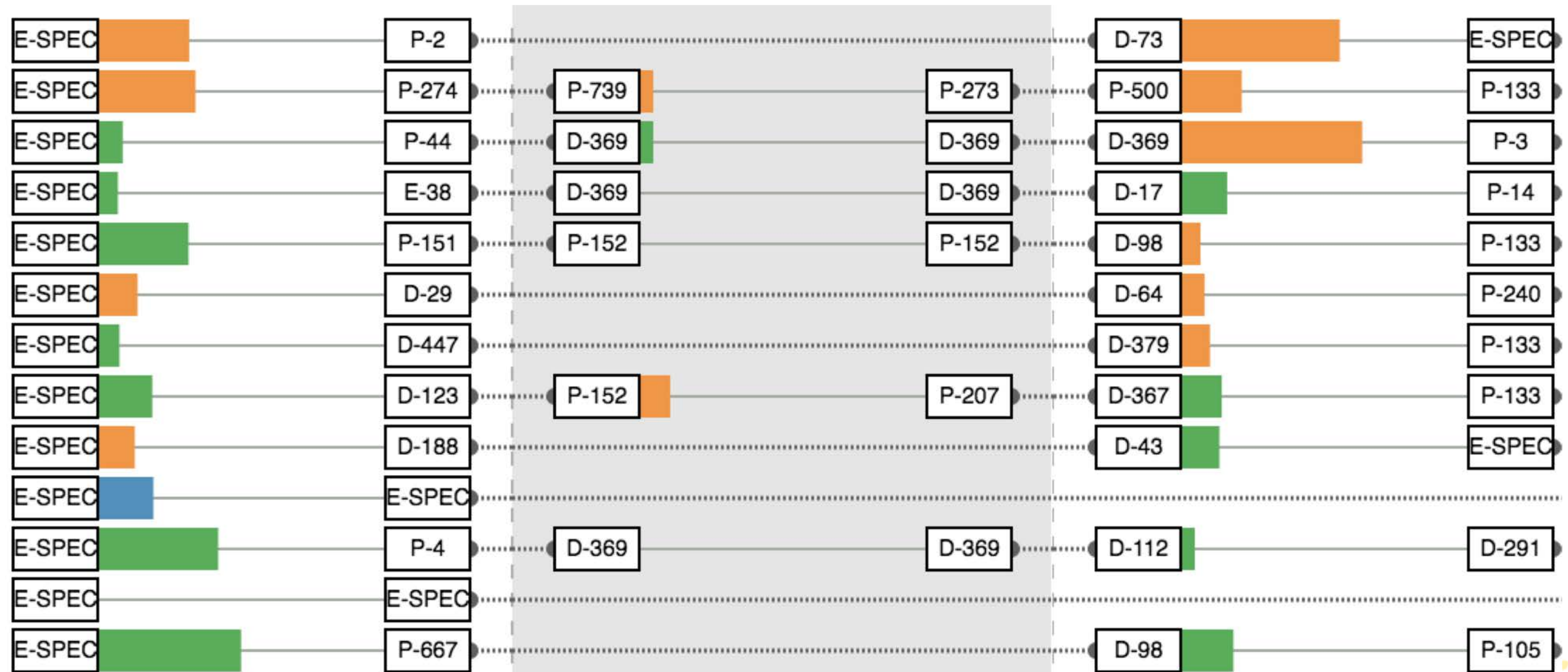




Sequence View



Sequence View



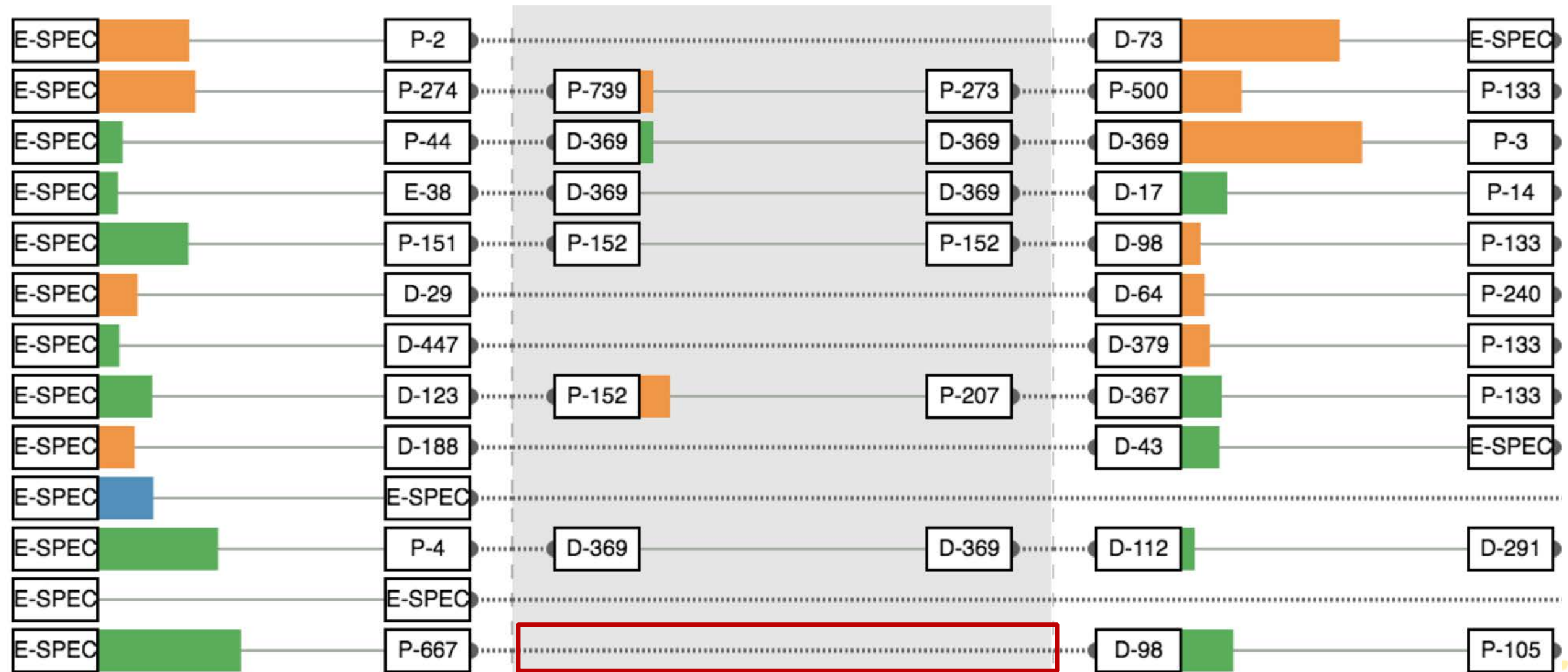
stage segment

D-98



P-105

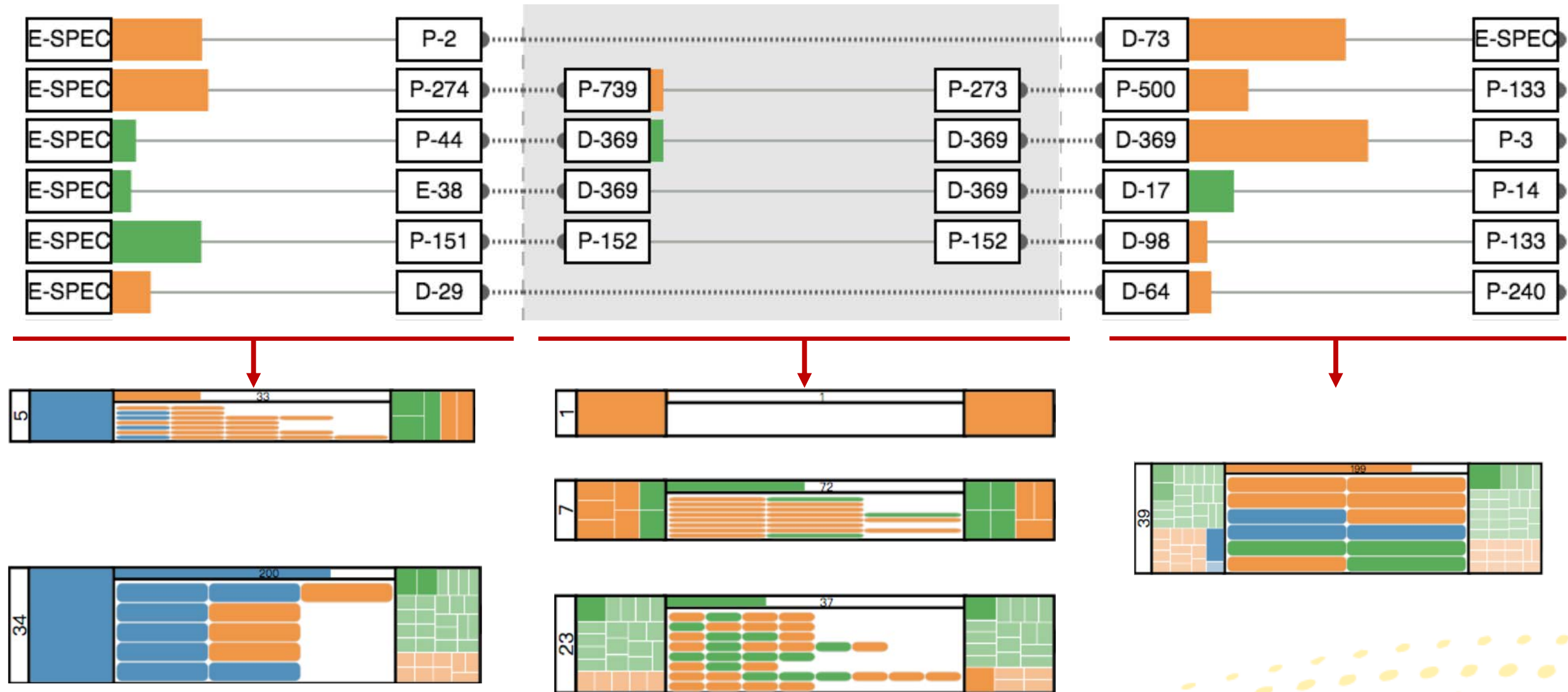
Sequence View



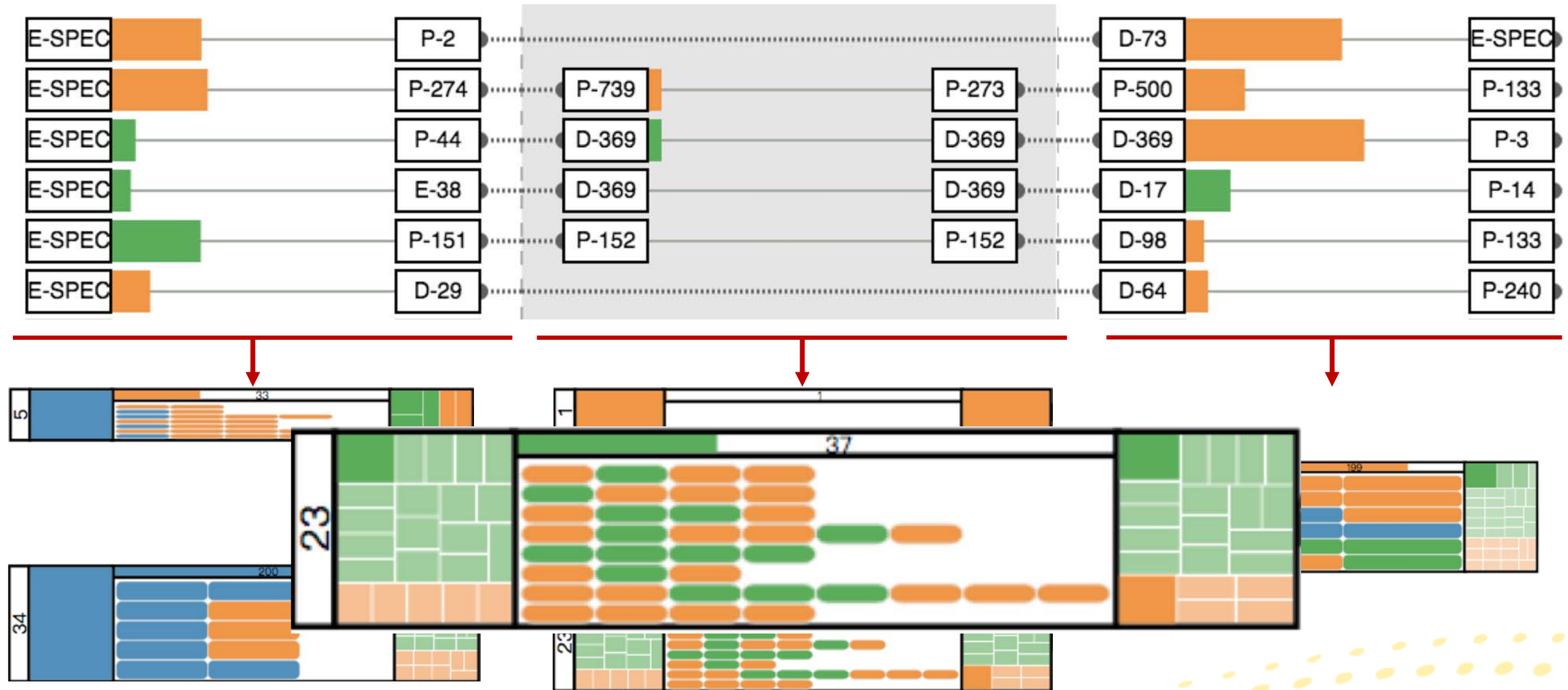
stage segment



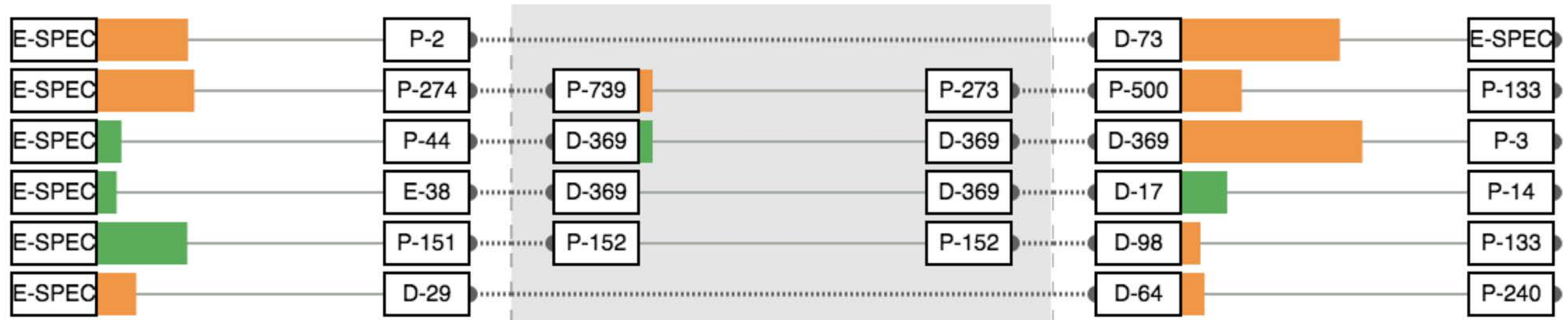
Cluster View



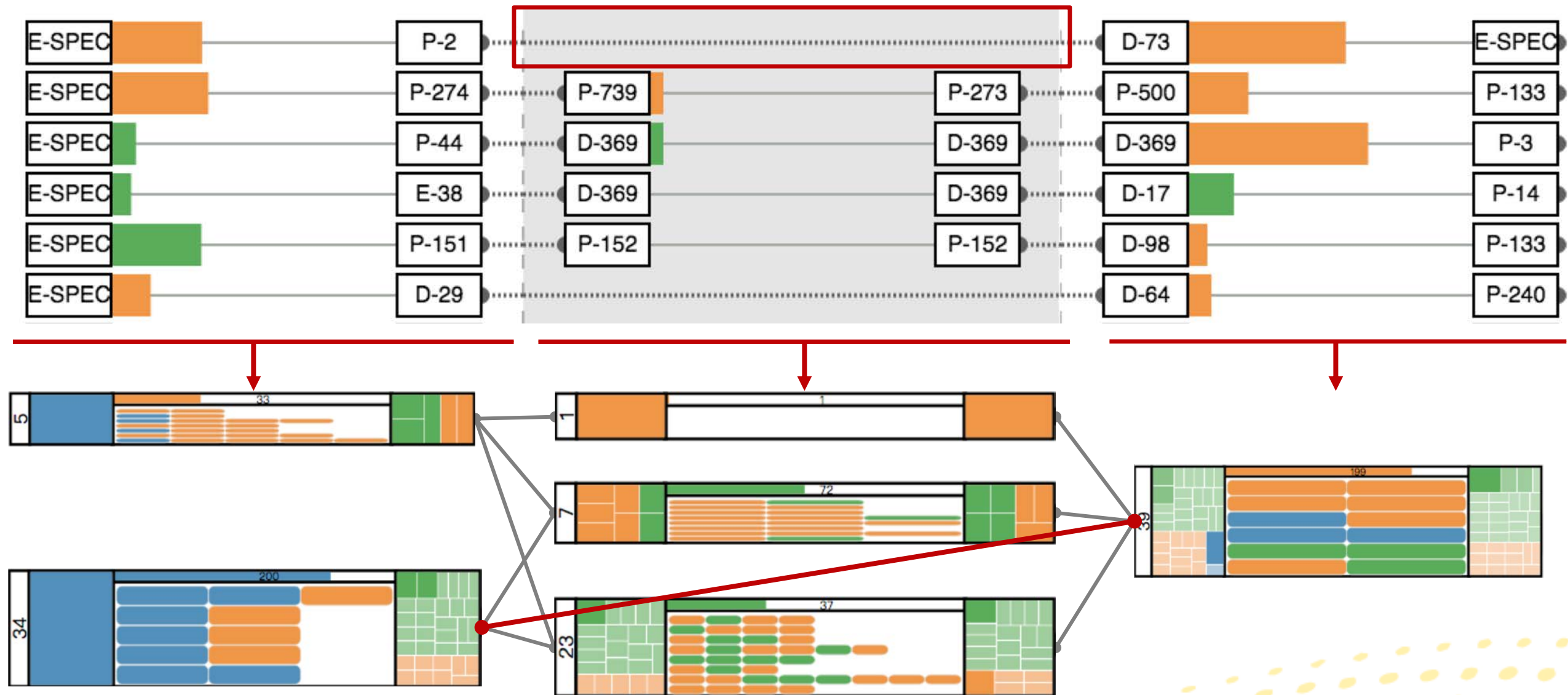
Cluster View



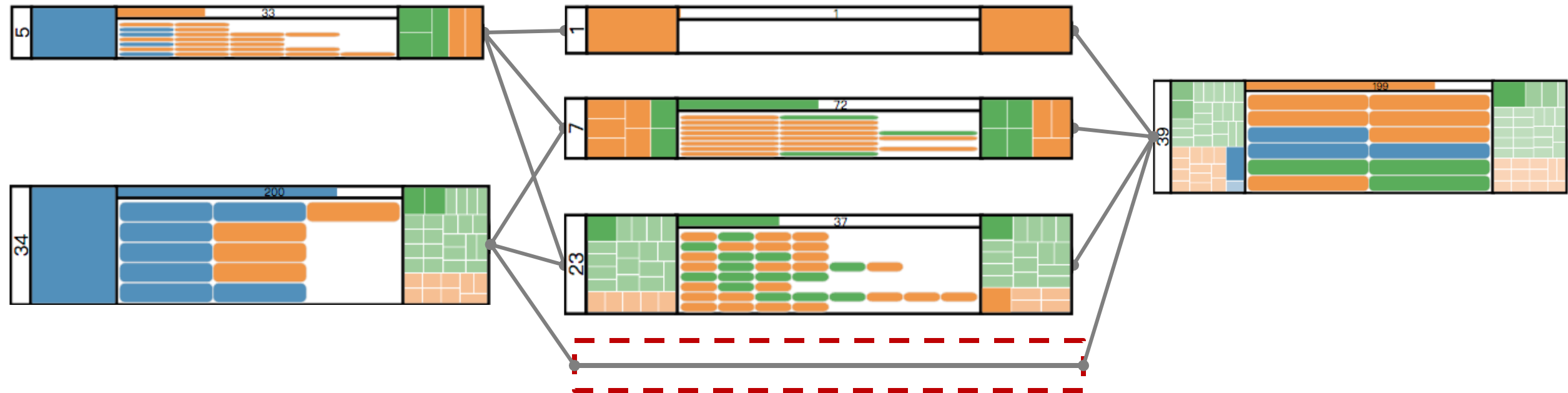
Cluster View



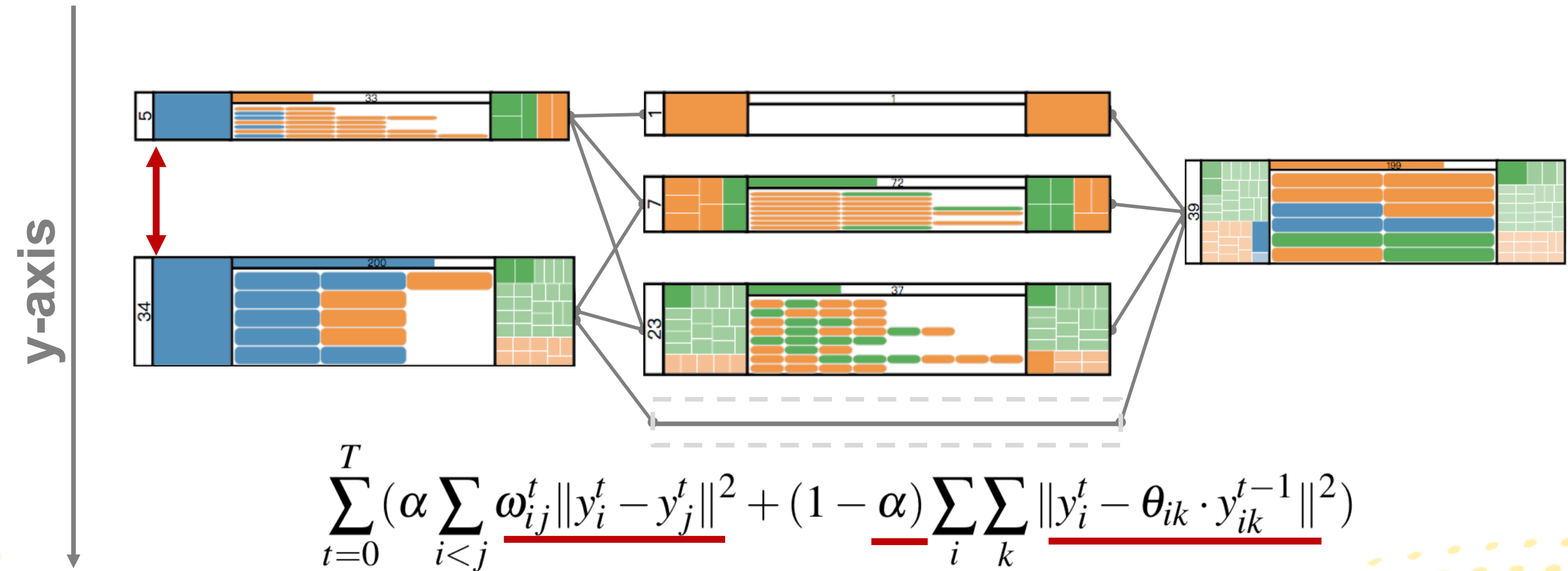
Layout Algorithm of Cluster View



Step#1 Insertion of the Intermediate Nodes



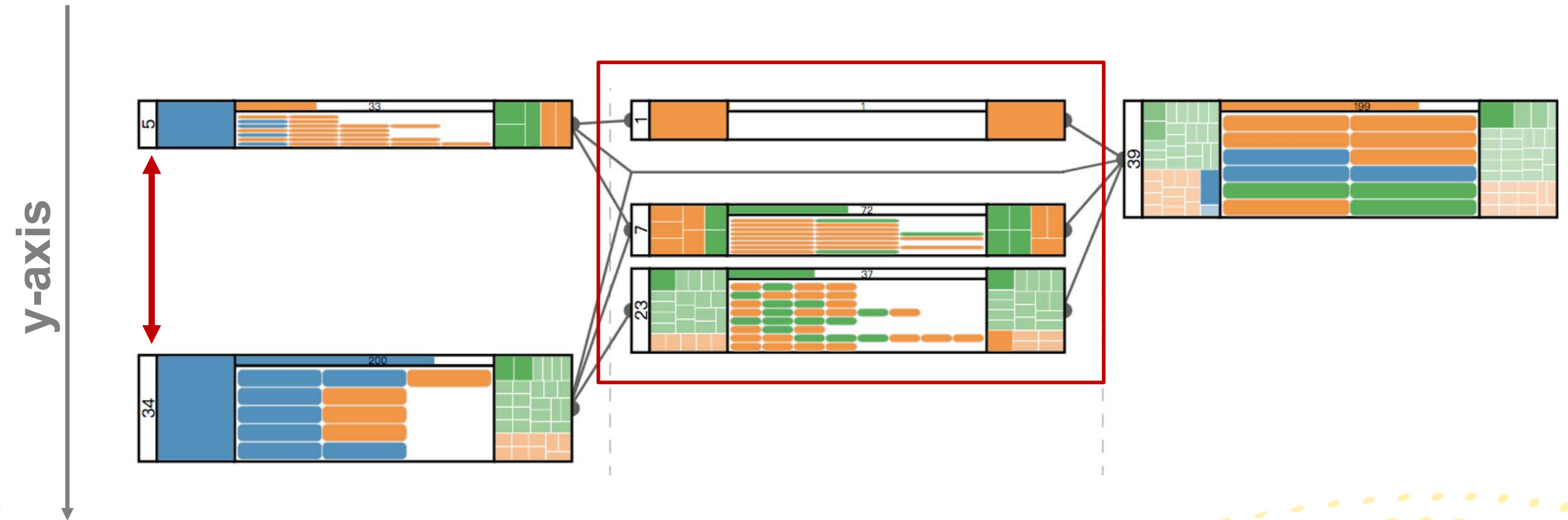
Step#2 Positioning of the Cluster Nodes

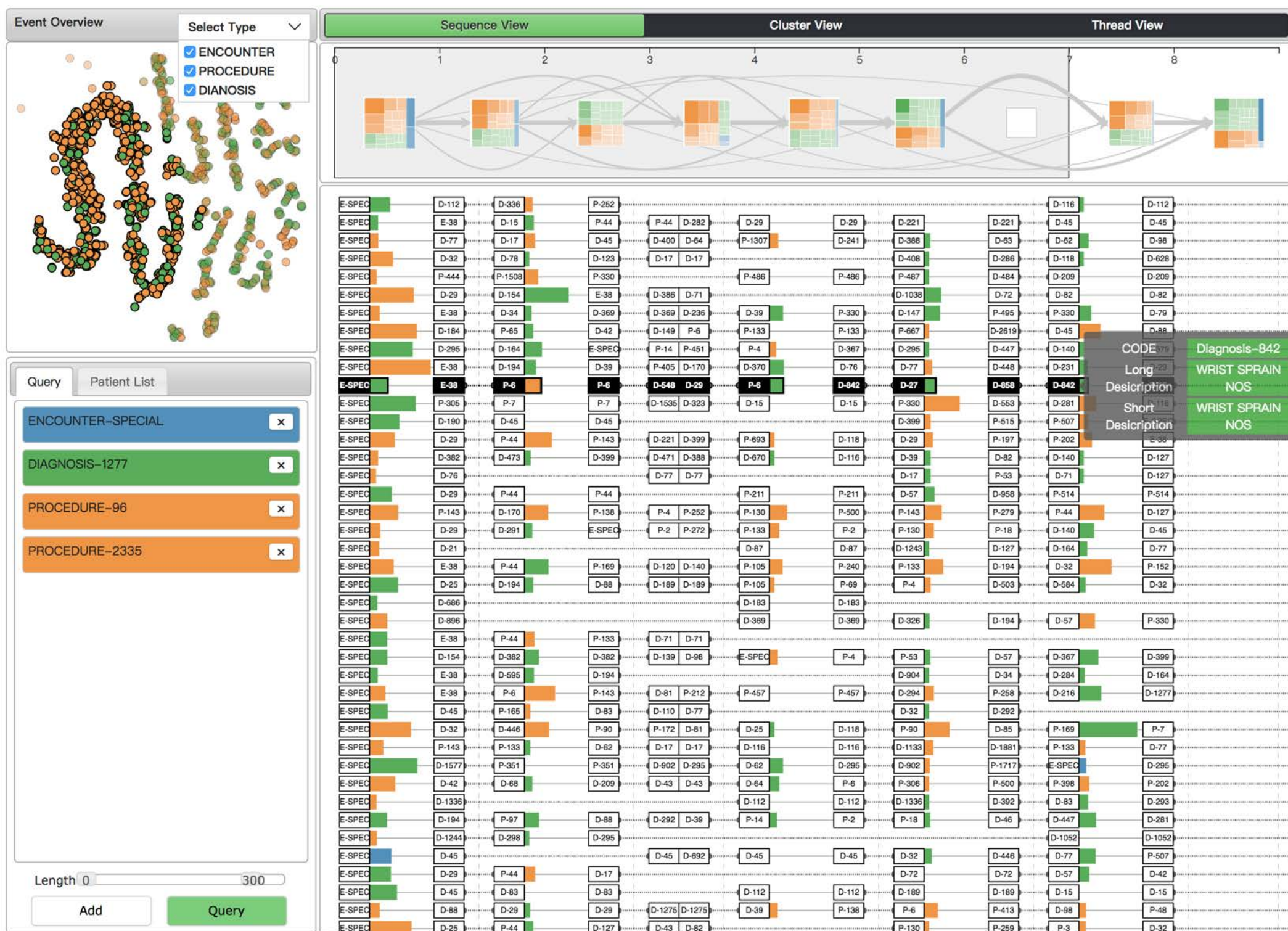


Group similar nodes

Minimize link crossing

Step#2 Positioning of the Cluster Nodes

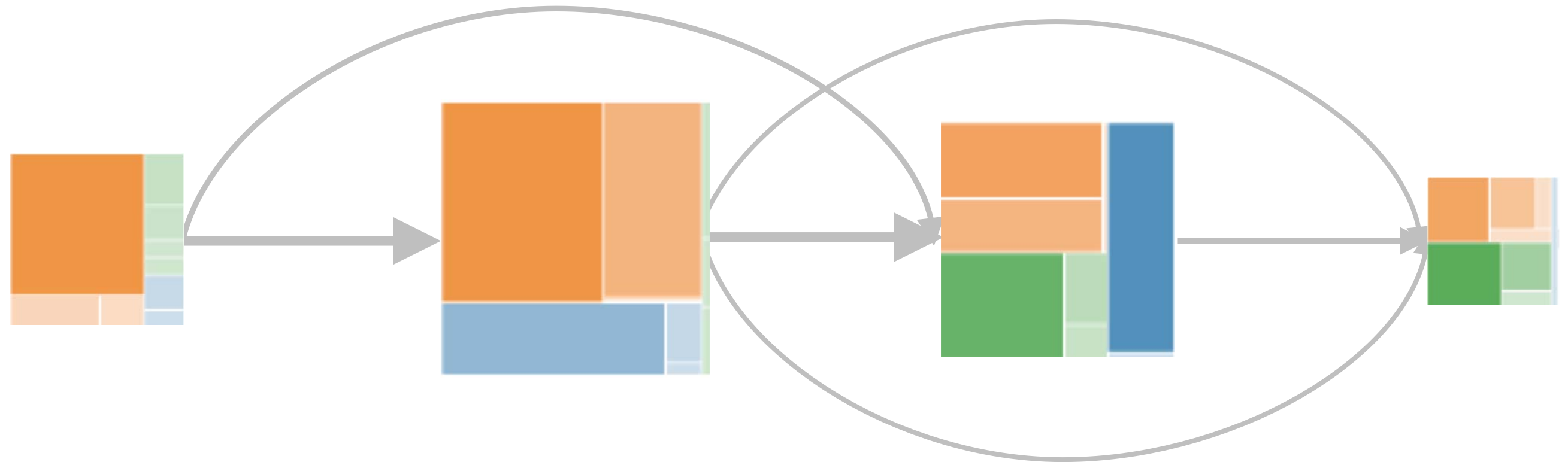




Summarization Level

- ✓ 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:

Add

Evaluation

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

Evaluation

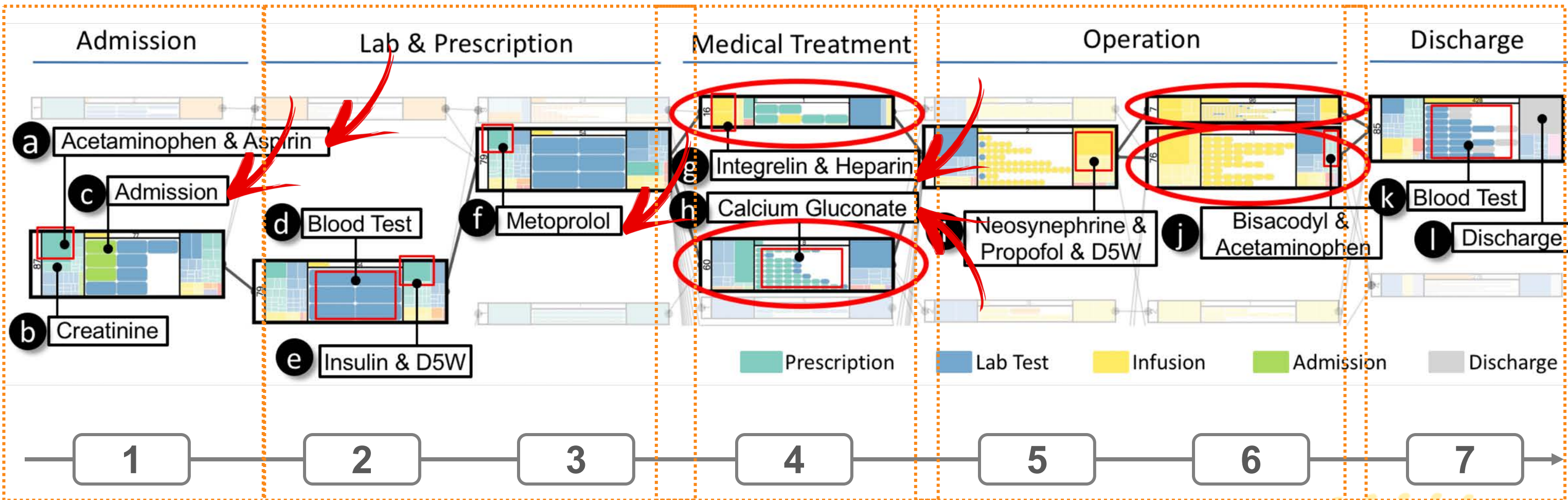
- Case studies on **MIMIC** and academic behavior datasets
- Expert interview
- Performance evaluation on **algorithm** and visualization

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.

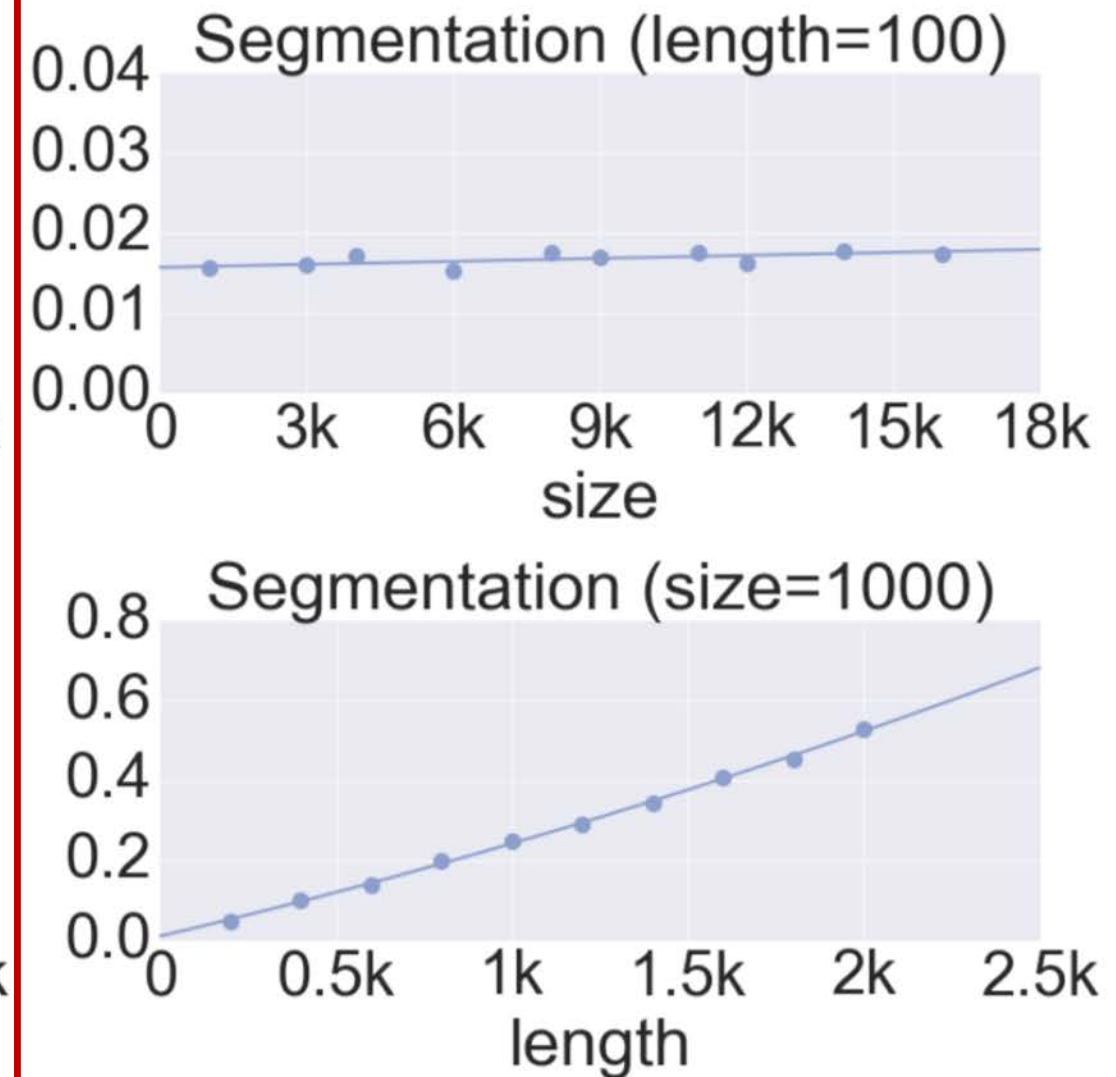
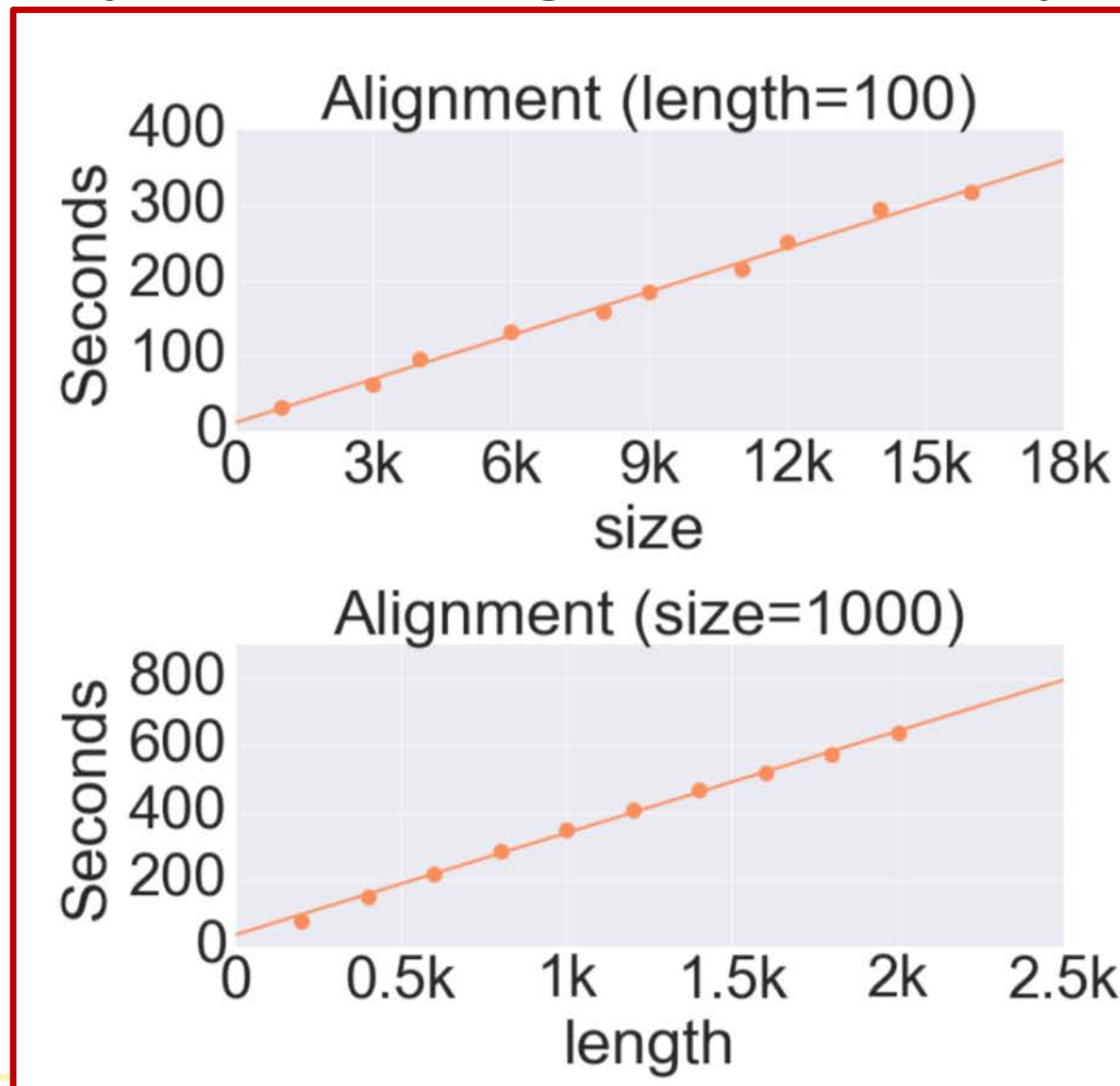


Case Study – MIMIC Dataset



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

