

Machine Learning for Complex Event Recognition

Nikos Katzouris¹, Alexander Artikis¹,
Georgios Paliouras¹

<http://cer.iit.demokritos.gr>

¹NCSR Demokritos, Athens, Greece

Team Members

Tenured Researcher:

- ▶ Georgios Paliouras

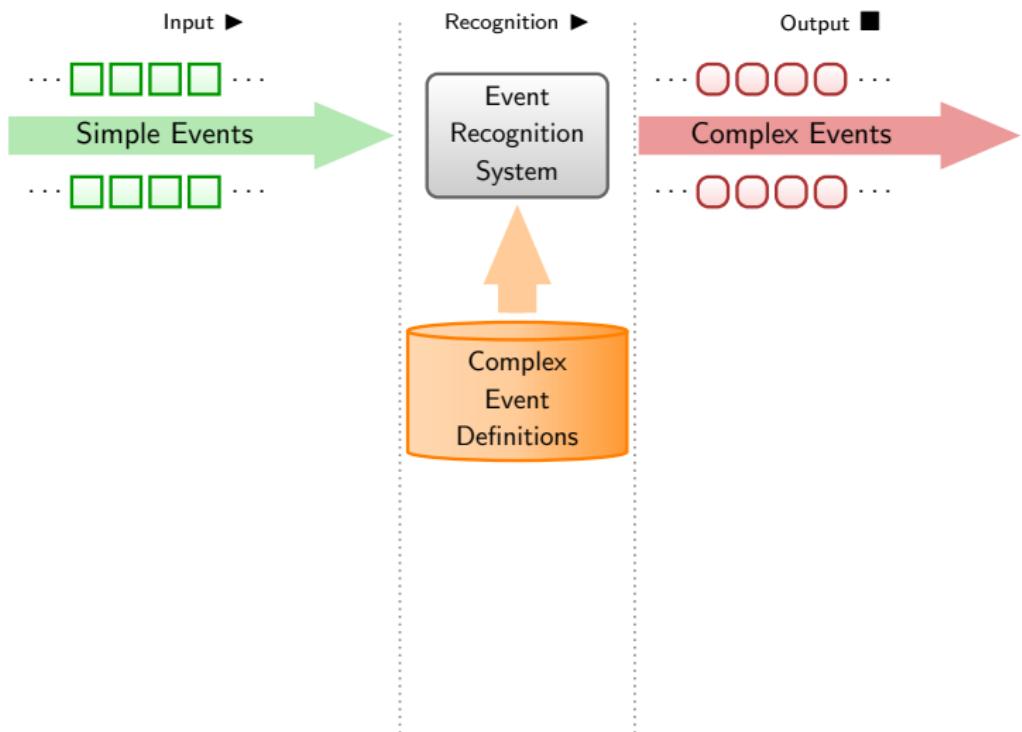
Research Associates:

- ▶ Elias Alevizos
- ▶ Alexander Artikis
- ▶ Nikos Katzouris
- ▶ Ioannis Kontopoulos
- ▶ Vagelis Michelioudakis
- ▶ Efthimis Tsilionis
- ▶ Christos Vlassopoulos

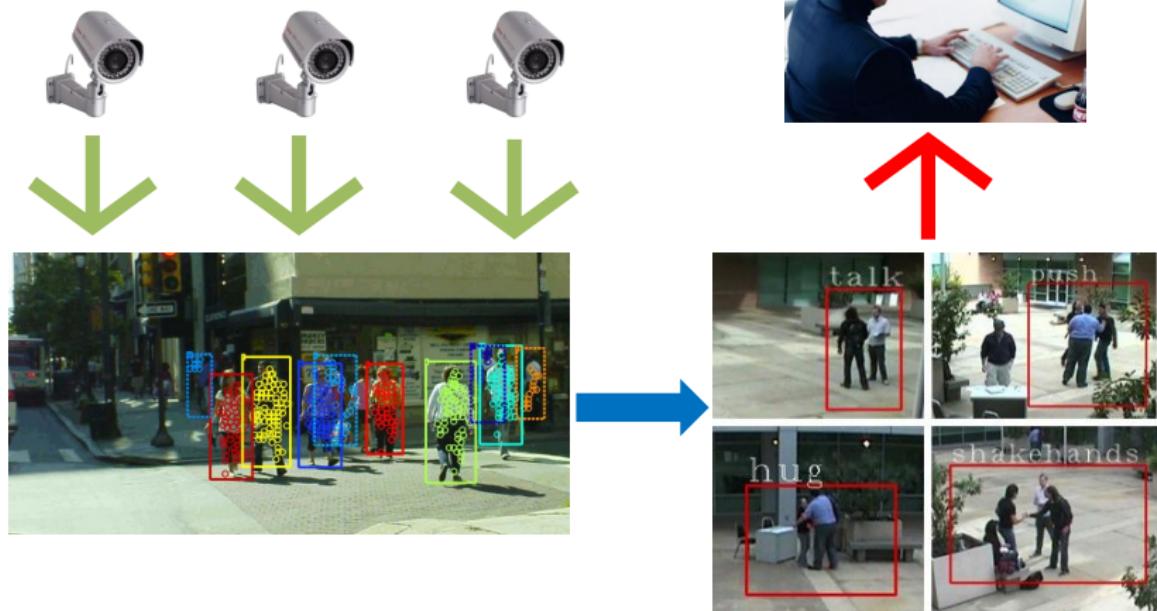
Website:

- ▶ <http://cer.iit.demokritos.gr>

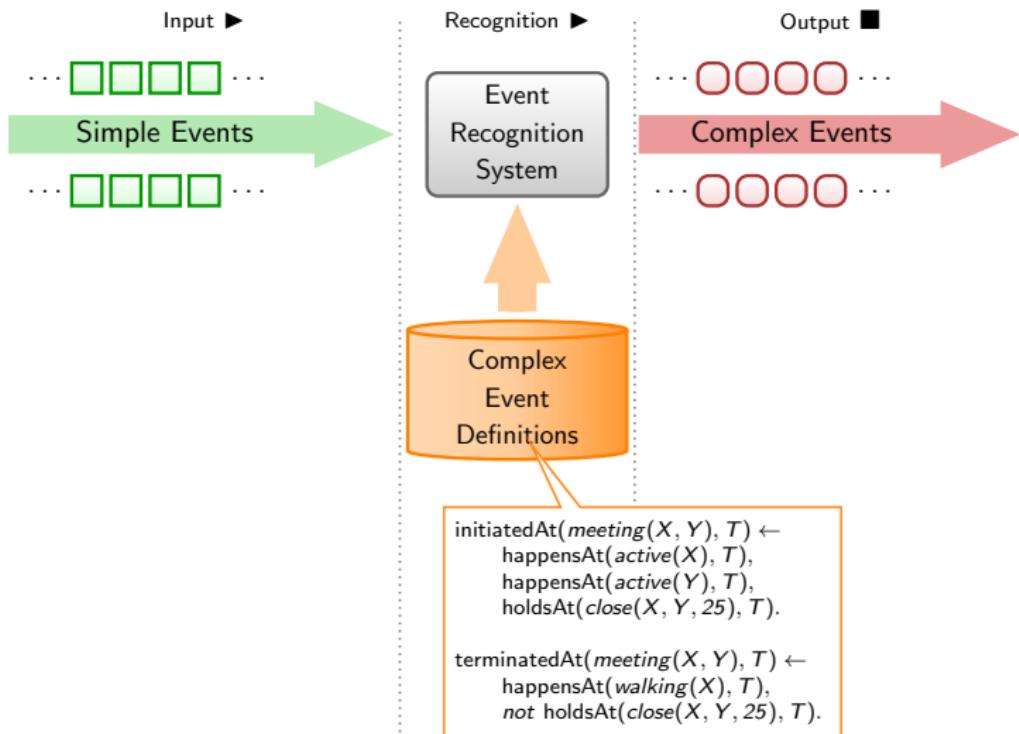
Complex Event Recognition



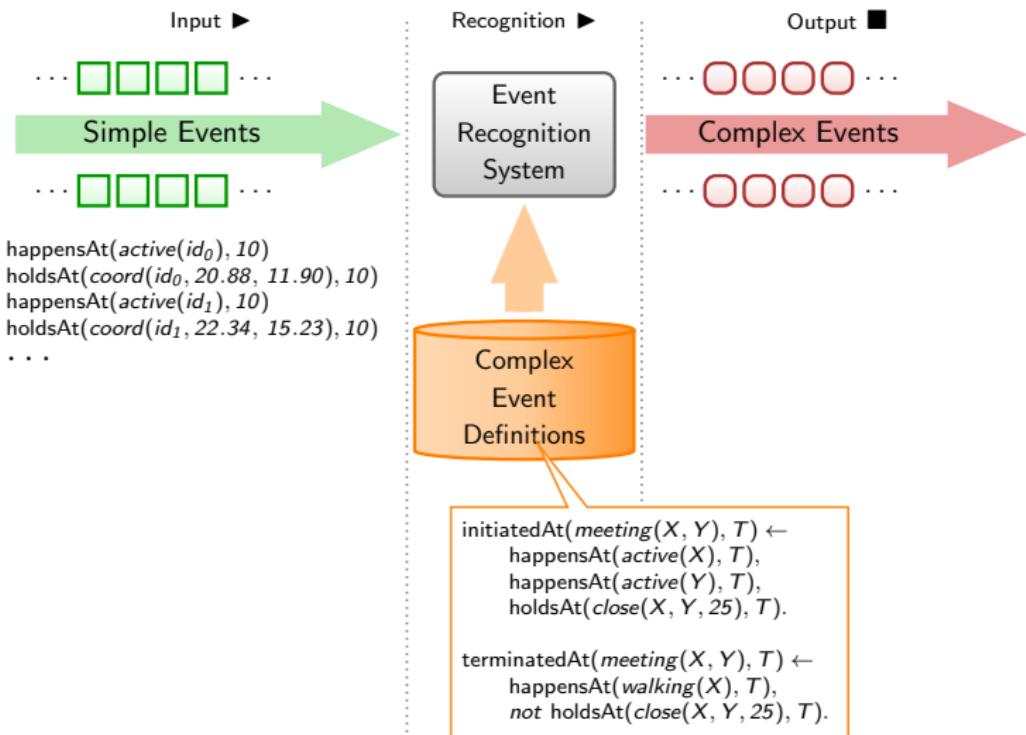
Complex Event Recognition



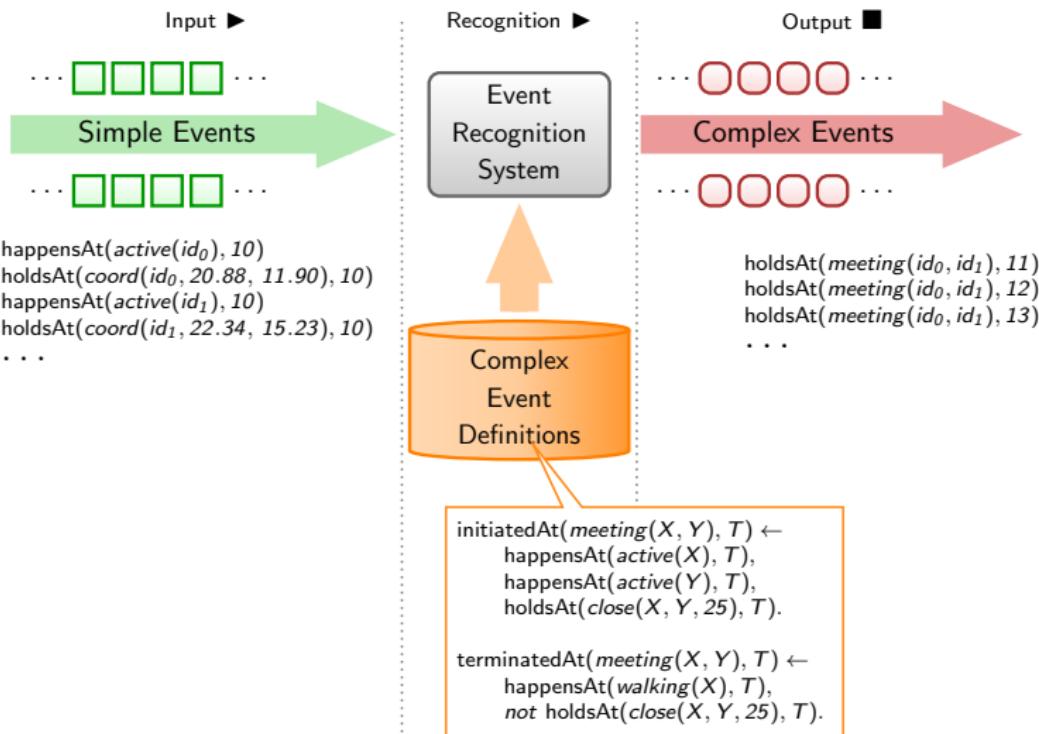
Complex Event Recognition



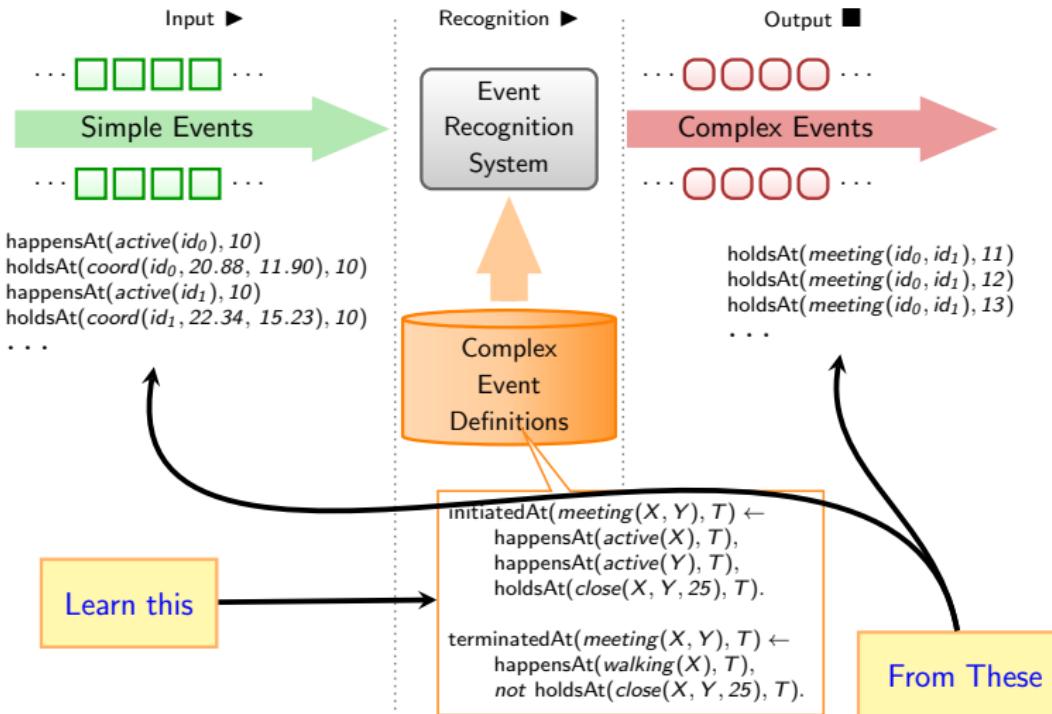
Complex Event Recognition



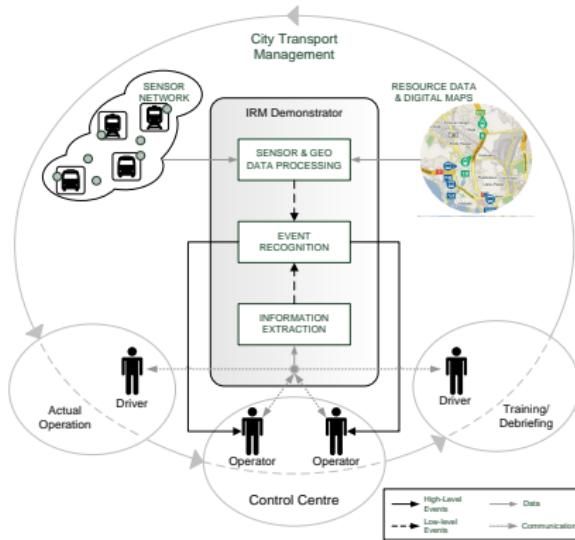
Complex Event Recognition



Learning for Complex Event Recognition



Applications



Complex Event Recognition using the Event Calculus

- ▶ Formal, declarative semantics.
- ▶ Representation of complex temporal phenomena.
- ▶ Representation of complex atemporal phenomena.
- ▶ Very efficient reasoning → RTEC.

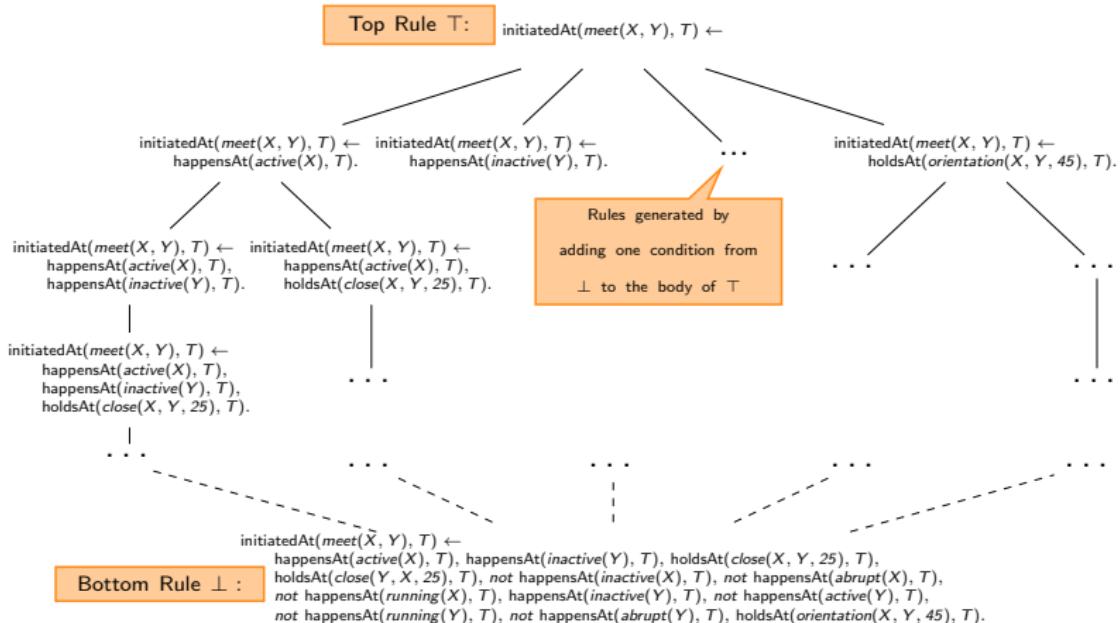
Complex Event Recognition using the Event Calculus

- ▶ Formal, declarative semantics.
- ▶ Representation of complex temporal phenomena.
- ▶ Representation of complex atemporal phenomena.
- ▶ Very efficient reasoning → RTEC.
- ▶ Direct connections to machine learning → Inductive Logic Programming (ILP).

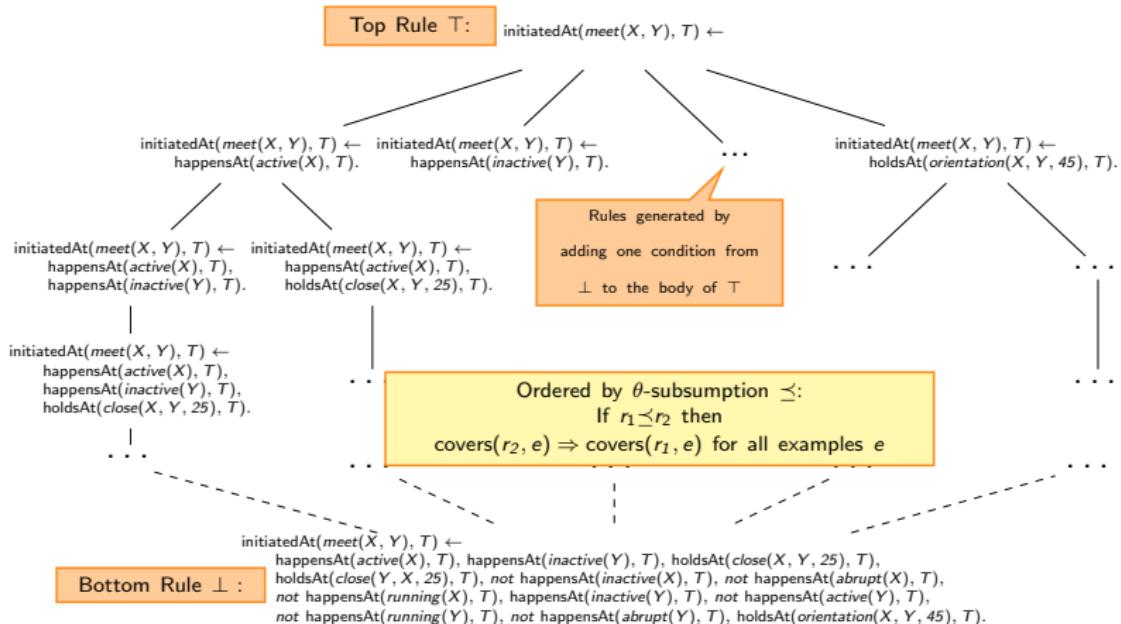
Inductive Logic Programming

- ▶ Input:
 - ▶ Positive and negative examples.
 - ▶ Background knowledge.
 - ▶ Language bias.
- ▶ Output:
 - ▶ A logical theory that entails as many positive and as few negative examples as possible.

Search Space for Rule Learning



Search Space for Rule Learning



Learning a Rule: Search

Positives
not covered

Top Rule T:

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt}(\text{active}(X), T).$

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt}(\text{inactive}(Y), T).$

...

...

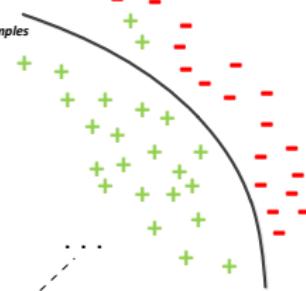
$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt}(\text{active}(X), T),$
 $\text{happensAt}(\text{inactive}(Y), T).$

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt}(\text{active}(X), T),$
 $\text{holdsAt}(\text{close}(X, Y, 25), T).$

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt}(\text{active}(X), T),$
 $\text{happensAt}(\text{inactive}(Y), T),$
 $\text{holdsAt}(\text{close}(X, Y, 25), T).$

...

Covered examples



Bottom Rule \perp :

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt}(\text{active}(X), T), \text{happensAt}(\text{inactive}(Y), T), \text{holdsAt}(\text{close}(X, Y, 25), T),$
 $\text{holdsAt}(\text{close}(Y, X, 25), T), \text{not happensAt}(\text{inactive}(X), T), \text{not happensAt}(\text{abrupt}(X), T),$
 $\text{not happensAt}(\text{running}(X), T), \text{happensAt}(\text{inactive}(Y), T), \text{not happensAt}(\text{active}(Y), T),$
 $\text{not happensAt}(\text{running}(Y), T), \text{not happensAt}(\text{abrupt}(Y), T), \text{holdsAt}(\text{orientation}(X, Y, 45), T).$

Learning a Rule: Search

Generalization ↑

Top Rule T:

initiatedAt(meet(X, Y), T) ←

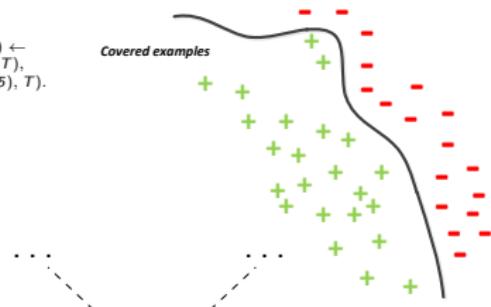
initiatedAt(meet(X, Y), T) ←
happensAt(active(X), T),
happensAt(inactive(Y), T).

initiatedAt(meet(X, Y), T) ←
happensAt(active(X), T),
happensAt(inactive(Y), T),
holdsAt(close(X, Y, 25), T).

Bottom Rule ⊥ :

initiatedAt(meet(X, Y), T) ←
happensAt(active(X), T), happensAt(inactive(Y), T), holdsAt(close(X, Y, 25), T),
holdsAt(close(Y, X, 25), T), not happensAt(inactive(X), T), not happensAt(abrupt(X), T),
not happensAt(running(X), T), happensAt(inactive(Y), T), not happensAt(active(Y), T),
not happensAt(running(Y), T), not happensAt(abrupt(Y), T), holdsAt(orientation(X, Y, 45), T).

Covered examples



Learning a Rule: Search

Negatives covered

Top Rule T:

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt(active}(X), T).$

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt(active}(X), T),$
 $\text{happensAt(inactive}(Y), T).$

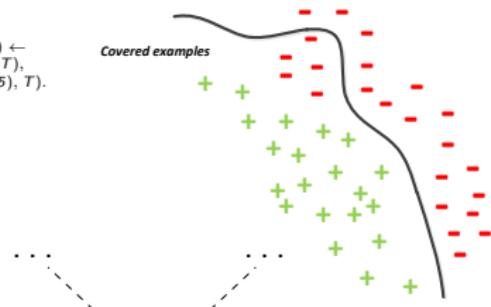
$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt(active}(X), T),$
 $\text{happensAt(inactive}(Y), T).$

$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt(active}(X), T),$
 $\text{happensAt(inactive}(Y), T),$
 $\text{holdsAt(close}(X, Y, 25), T).$

Bottom Rule ⊥ :

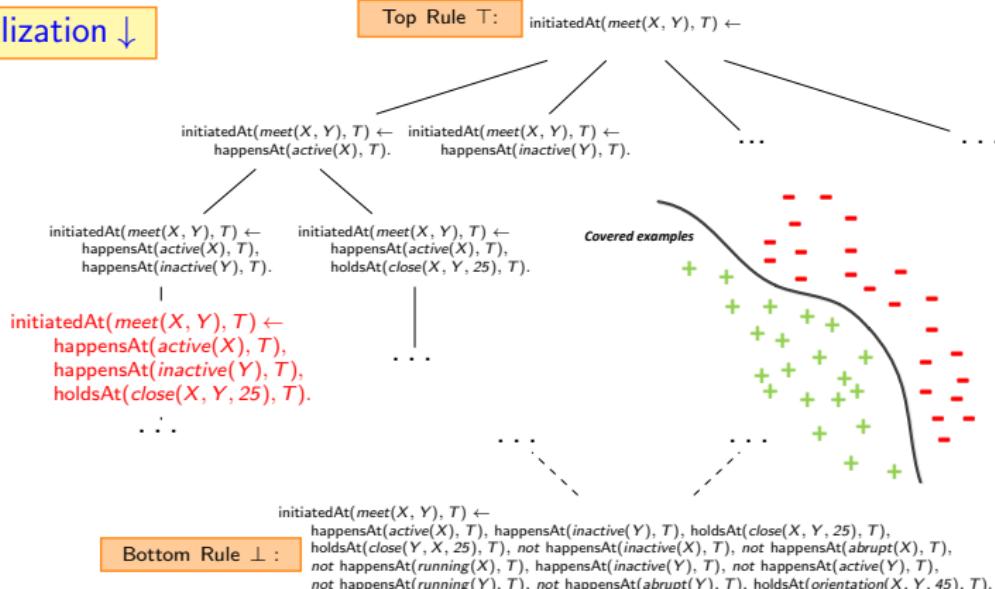
$\text{initiatedAt(meet}(X, Y), T) \leftarrow$
 $\text{happensAt(active}(X), T), \text{happensAt(inactive}(Y), T), \text{holdsAt(close}(X, Y, 25), T),$
 $\text{holdsAt(close}(Y, X, 25), T), \text{not happensAt(inactive}(X), T), \text{not happensAt(abrupt}(X), T),$
 $\text{not happensAt(running}(X), T), \text{happensAt(inactive}(Y), T), \text{not happensAt(active}(Y), T),$
 $\text{not happensAt(running}(Y), T), \text{not happensAt(abrupt}(Y), T), \text{holdsAt(orientation}(X, Y, 45), T).$

Covered examples



Learning a Rule: Search

Specialization ↓



Online Inductive Logic Programming

Challenge:

- ▶ Inductive Logic Programming algorithms are batch learners.
 - ▶ Each candidate in the search space is evaluated **on the entire dataset.**

Online Inductive Logic Programming

Challenge:

- ▶ Inductive Logic Programming algorithms are batch learners.
 - ▶ Each candidate in the search space is evaluated **on the entire dataset.**

Goal:

- ▶ Online learning:
 - ▶ Examples arrive in a stream.
 - ▶ Each example is “seen” once.

Online Inductive Logic Programming

Challenge:

- ▶ Inductive Logic Programming algorithms are batch learners.
 - ▶ Each candidate in the search space is evaluated **on the entire dataset**.

Goal:

- ▶ Online learning:
 - ▶ Examples arrive in a stream.
 - ▶ Each example is “seen” once.

Approach:

- ▶ Make decisions from subsets of the stream:
 - ▶ Decisions are optimal “locally”.
 - ▶ Decisions are optimal “globally” ...
 - ▶ within an error margin ϵ ,
 - ▶ with probability $1-\delta$.

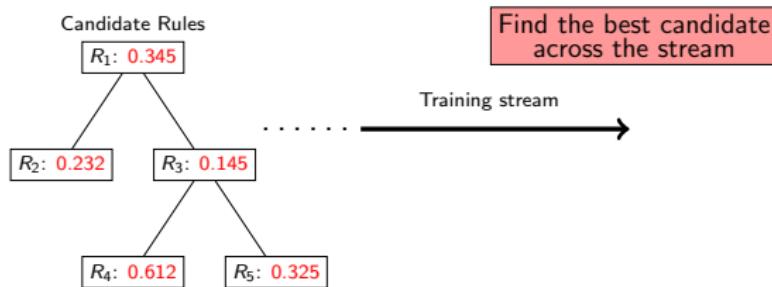
The Hoeffding Bound

- ▶ X is a random variable.
- ▶ X_1, \dots, X_N are N independent observations of X 's values.
- ▶ Let \bar{X} be the known, **observed mean** of X .
- ▶ Let \hat{X} be the unknown, **true mean** of X .

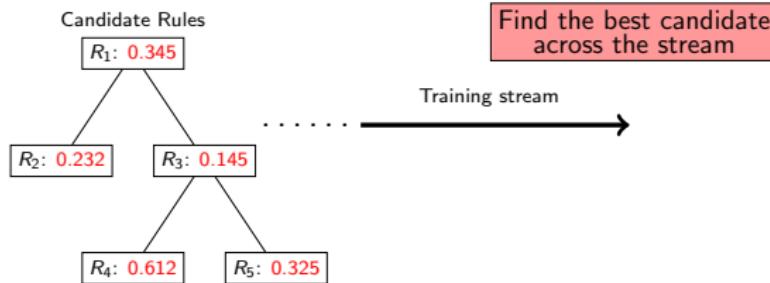
The Hoeffding Bound

- ▶ X is a random variable.
- ▶ X_1, \dots, X_N are N independent observations of X 's values.
- ▶ Let \bar{X} be the known, **observed mean** of X .
- ▶ Let \hat{X} be the unknown, **true mean** of X .
- ▶ Then:
$$\bar{X} - \epsilon \leq \hat{X} \leq \bar{X} + \epsilon, \text{ with probability } 1 - \delta, \text{ where } \epsilon = \sqrt{\frac{\ln(1/\delta)}{2N}}$$

Online Rule Learning



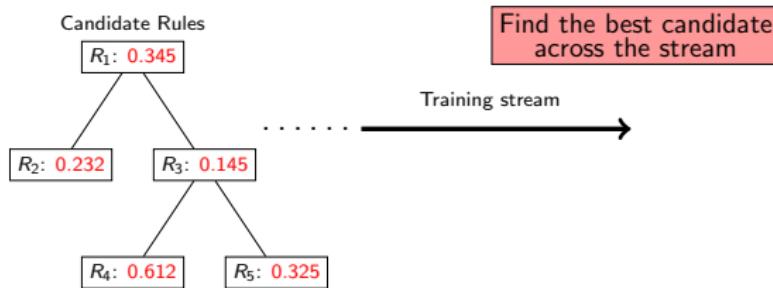
Online Rule Learning



As examples stream in...

$$\text{Monitor } \bar{X} = \overline{\text{score}}_{\text{BestRule}} - \overline{\text{score}}_{\text{SecondBestRule}}$$

Online Rule Learning



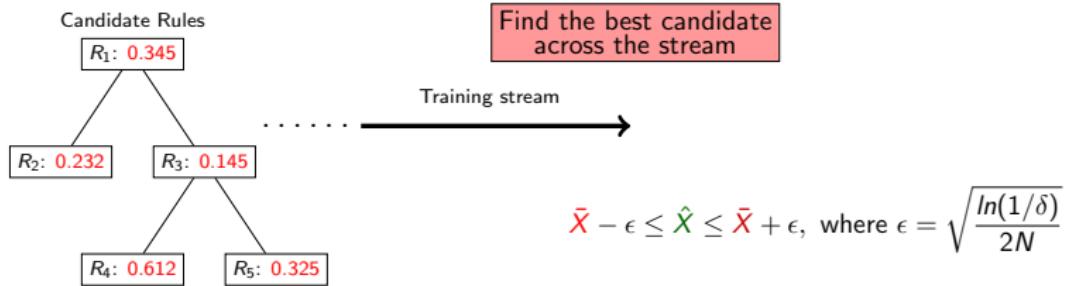
As examples stream in...

$$\text{Monitor } \bar{X} = \overline{\text{score}}_{\text{BestRule}} - \overline{\text{score}}_{\text{SecondBestRule}}$$

Continue until the number N of examples

$$\text{makes } \bar{X} > \epsilon = \sqrt{\frac{\ln(1/\delta)}{2N}}$$

Online Rule Learning



As examples stream in...

$$\text{Monitor } \bar{X} = \overline{\text{score}}_{\text{BestRule}} - \overline{\text{score}}_{\text{SecondBestRule}}$$

Continue until the number N of examples

$$\text{makes } \bar{X} > \epsilon = \sqrt{\frac{\ln(1/\delta)}{2N}}$$

Then

$$\bar{X} - \epsilon > 0 \Rightarrow$$

$$\hat{X} > 0 \Rightarrow$$

BestRule is indeed the best rule,
with probability $1-\delta$.

Online Rule Learning

initiatedAt(*meet(X, Y)*, *T*)
score: 0.987

initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(Y)*, *T*). happensAt(*active(X)*, *T*). holdsAt(*close(X, Y, 25)*, *T*).
score: 0.312 score: 0.534 score: 0.023

Input stream



Online Rule Learning

initiatedAt(*meet(X, Y)*, *T*)
score: 0.732

initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(Y)*, *T*). happensAt(*active(X)*, *T*). holdsAt(*close(X, Y, 25)*, *T*).
score: 0.307 score: 0.568 score: 0.048

Input stream



Online Rule Learning

initiatedAt(*meet(X, Y)*, *T*)
score: 0.216

initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(Y)*, *T*). happensAt(*active(X)*, *T*). holdsAt(*close(X, Y, 25)*, *T*).
score: 0.418 score: 0.632 score: 0.122

Input stream



Online Rule Learning

initiatedAt(*meet(X, Y)*, *T*)

initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(Y)*, *T*).

initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(X)*, *T*).

initiatedAt(*meet(X, Y)*, *T*) \leftarrow
holdsAt(*close(X, Y, 25)*, *T*).

Best rule so far

Used $\mathcal{O}\left(\frac{1}{\epsilon^2} \ln \frac{1}{\delta}\right)$ examples

Online Rule Learning

initiatedAt(*meet*(*X*, *Y*), *T*)

initiatedAt(*meet*(*X*, *Y*), *T*) \leftarrow happensAt(*active*(*Y*), *T*). initiatedAt(*meet*(*X*, *Y*), *T*) \leftarrow happensAt(*active*(*X*), *T*). initiatedAt(*meet*(*X*, *Y*), *T*) \leftarrow holdsAt(*close*(*X*, *Y*, 25), *T*).
score: 0.644

initiatedAt(*meet*(*X*, *Y*), *T*) \leftarrow happensAt(*active*(*Y*), *T*), initiatedAt(*meet*(*X*, *Y*), *T*) \leftarrow happensAt(*active*(*Y*), *T*),
happensAt(*active*(*X*), *T*). holdsAt(*close*(*X*, *Y*, 25), *T*).
score: 0.618 score: 0.588

Input stream



Online Rule Learning

`initiatedAt(meet(X, Y), T)`

`initiatedAt(meet(X, Y), T) ←
happensAt(active(Y), T).`

`score: 0.689`

`initiatedAt(meet(X, Y), T) ←
happensAt(active(X), T).`

`initiatedAt(meet(X, Y), T) ←
happensAt(active(Y), T),
happensAt(active(X), T).`

`score: 0.656`

`initiatedAt(meet(X, Y), T) ←
happensAt(active(Y), T),
holdsAt(close(X, Y, 25), T).`

`score: 0.522`

Input stream



Online Rule Learning

initiatedAt(*meet(X, Y)*, *T*)

initiatedAt(*meet(X, Y)*, *T*) \leftarrow happensAt(*active(Y)*, *T*). initiatedAt(*meet(X, Y)*, *T*) \leftarrow happensAt(*active(X)*, *T*). initiatedAt(*meet(X, Y)*, *T*) \leftarrow holdsAt(*close(X, Y, 25)*, *T*).
score: 0.672

initiatedAt(*meet(X, Y)*, *T*) \leftarrow happensAt(*active(Y)*, *T*), initiatedAt(*meet(X, Y)*, *T*) \leftarrow happensAt(*active(Y)*, *T*),
happensAt(*active(X)*, *T*). holdsAt(*close(X, Y, 25)*, *T*).
score: 0.693 score: 0.503

Input stream



Online Rule Learning

initiatedAt(*meet(X, Y)*, *T*)

initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(Y)*, *T*). happensAt(*active(X)*, *T*). holdsAt(*close(X, Y, 25)*, *T*).

initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(Y)*, *T*), happensAt(*active(Y)*, *T*),
happensAt(*active(X)*, *T*). holdsAt(*close(X, Y, 25)*, *T*).

Best rule so far

Used $\mathcal{O}\left(\frac{1}{\epsilon^2} \ln \frac{1}{\delta}\right)$ examples

Online Rule Learning

initiatedAt(*meet(X, Y)*, *T*)

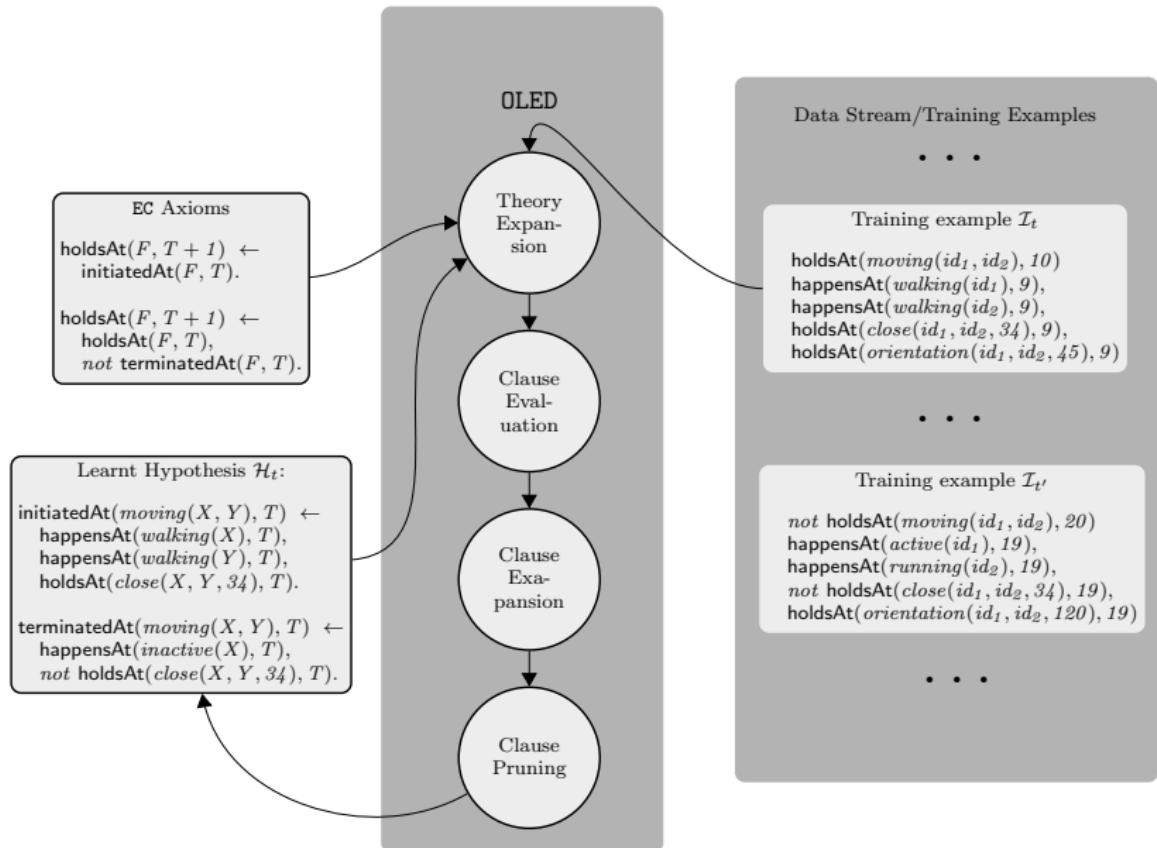
initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(Y)*, *T*). happensAt(*active(X)*, *T*). holdsAt(*close(X, Y, 25)*, *T*).

initiatedAt(*meet(X, Y)*, *T*) \leftarrow initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(Y)*, *T*), happensAt(*active(Y)*, *T*),
happensAt(*active(X)*, *T*). holdsAt(*close(X, Y, 25)*, *T*).
score: 0.693

initiatedAt(*meet(X, Y)*, *T*) \leftarrow
happensAt(*active(Y)*, *T*),
happensAt(*active(X)*, *T*),
holdsAt(*close(X, Y, 25)*, *T*).
score: 0.618

And so on...

Theory (Event Pattern Set) Learning



Empirical Evaluation

- ▶ Activity recognition using a benchmark dataset (CAVIAR).
 - ▶ 28 surveillance videos.
- ▶ Input: short-term activities per video frame+contextual information:
 - ▶ walking, active, inactive, running.
 - ▶ coordinates, orientation, occlusion.
- ▶ Learn concepts for *Move* and *Meet*.
- ▶ 10-fold cross-validation.

Empirical Evaluation

	Method	Precision	Recall	F₁-score	Time (sec)
<i>Move</i>	EC _{crisp}	0.909	0.634	0.751	–
	EC _{MM}	0.844	0.941	0.890	1692
	XHAIL	0.779	0.914	0.841	7836
	OLED	0.709	0.948	0.812	12
	OSL	–	–	–	N/A (> 25h)
	OSL α	0.7823	0.8882	0.8319	1342
<i>Meet</i>	EC _{crisp}	0.687	0.855	0.762	–
	EC _{MM}	0.919	0.813	0.863	1133
	XHAIL	0.804	0.927	0.861	7248
	OLED	0.943	0.750	0.836	23
	OSL	–	–	–	N/A (> 25h)
	OSL α	0.9055	0.8966	0.9010	174

Current work

- ▶ Statistical Relational Learning
 - ▶ Logic + Probability
 - ▶ Online structure + parameter learning.
- ▶ Distributed/parallel learning.

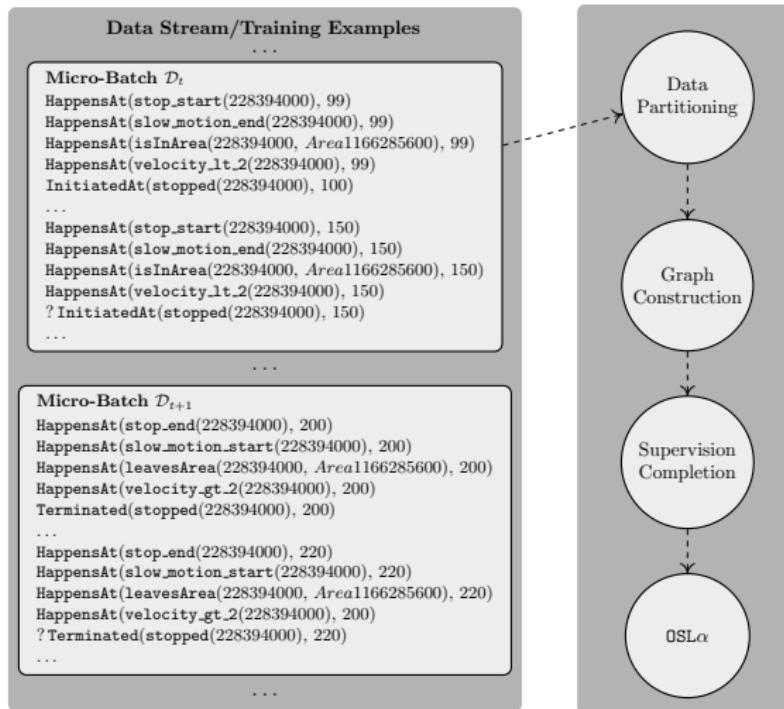
Resources

- ▶ N. Katzouris, A. Artikis and G. Paliouras. Online Learning of Event Definitions, TPLP 16(5-6), pp. 817-833, 2016
- ▶ N. Katzouris, A. Artikis and G. Paliouras. Parallel Online Learning of Event Definitions, ILP, 2017
- ▶ <http://github.com/nkatzz/OLED>

Semi-Supervised Online Structure Learning

- ▶ Common **problem** in Big data applications:
 - ▶ Incomplete annotation of activities of ‘special significance’.
 - ▶ Providing complete annotation is time-consuming.
 - ▶ The process of complex event pattern construction is hindered.
- ▶ **Solution:** Online supervision completion
 - ▶ Use given labels to complete the missing ones as data arrive.
 - ▶ Perform supervised structure learning.

Semi-Supervised Online Structure Learning



Data Partitioning

Training Sequence

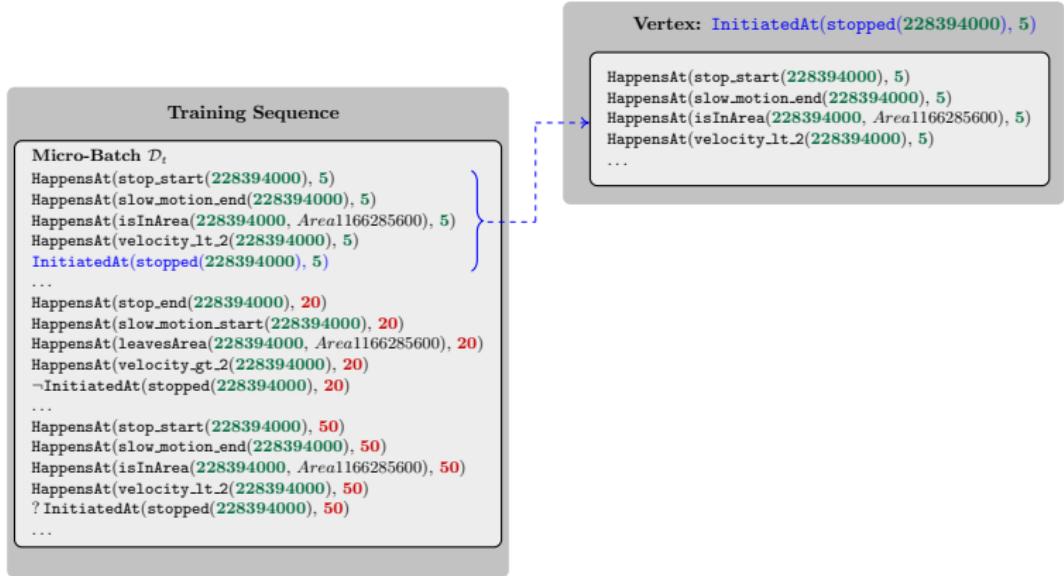
```
Micro-Batch  $D_t$ 
HappensAt(stop_start(228394000), 5)
HappensAt(slow_motion_end(228394000), 5)
HappensAt(isInArea(228394000, Area1166285600), 5)
HappensAt(velocity_lt_2(228394000), 5)
InitiatedAt(stopped(228394000), 5)
...
HappensAt(stop_end(228394000), 20)
HappensAt(slow_motion_start(228394000), 20)
HappensAt(leavesArea(228394000, Area1166285600), 20)
HappensAt(velocity_gt_2(228394000), 20)
~InitiatedAt(stopped(228394000), 20)
...
HappensAt(stop_start(228394000), 50)
HappensAt(slow_motion_end(228394000), 50)
HappensAt(isInArea(228394000, Area1166285600), 50)
HappensAt(velocity_lt_2(228394000), 50)
? InitiatedAt(stopped(228394000), 50)
...
```

Data Partitioning

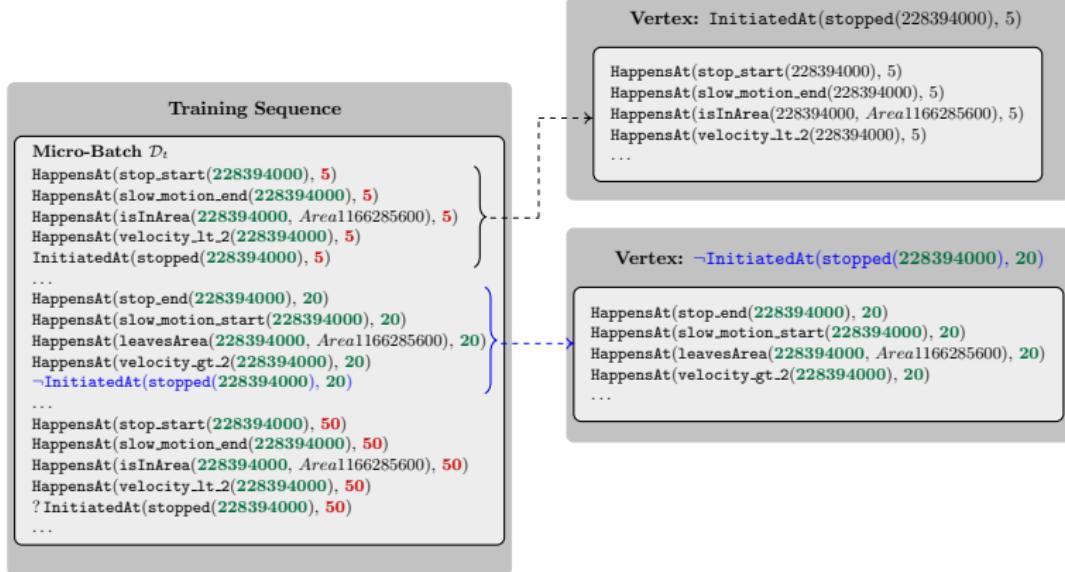
Training Sequence

```
Micro-Batch  $\mathcal{D}_t$ 
HappensAt(stop.start(228394000), 5)
HappensAt(slow_motion.end(228394000), 5)
HappensAt(isInArea(228394000, Area1166285600), 5)
HappensAt(velocity.lt_2(228394000), 5)
InitiatedAt(stopped(228394000), 5)
...
HappensAt(stop.end(228394000), 20)
HappensAt(slow_motion.start(228394000), 20)
HappensAt(leavesArea(228394000, Area1166285600), 20)
HappensAt(velocity.gt_2(228394000), 20)
¬InitiatedAt(stopped(228394000), 20)
...
HappensAt(stop.start(228394000), 50)
HappensAt(slow_motion.end(228394000), 50)
HappensAt(isInArea(228394000, Area1166285600), 50)
HappensAt(velocity.lt_2(228394000), 50)
? InitiatedAt(stopped(228394000), 50)
...
```

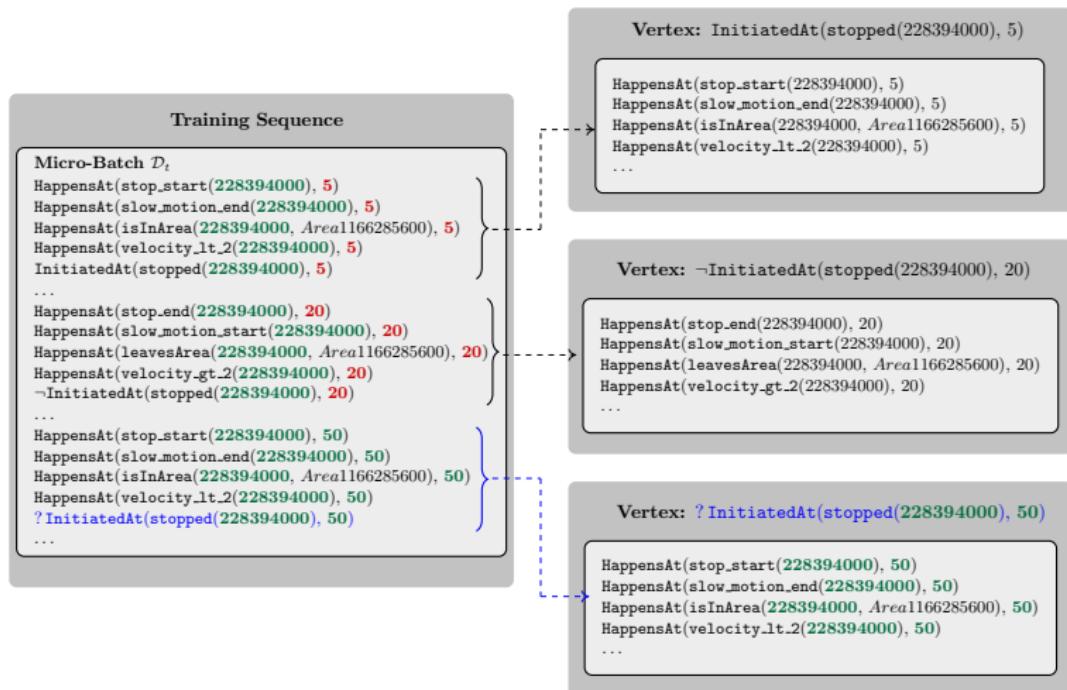
Data Partitioning



Data Partitioning



Data Partitioning



Graph Construction

Vertex: InitiatedAt(stopped(228394000), 5)

```
HappensAt(stop_start(228394000), 5)
HappensAt(slow_motion_end(228394000), 5)
HappensAt(isInArea(228394000, Area1166285600), 5)
HappensAt(velocity_lt_2(228394000), 5)
...
```

Vertex: \neg InitiatedAt(stopped(228394000), 20)

```
HappensAt(stop_end(228394000), 20)
HappensAt(slow_motion_start(228394000), 20)
HappensAt(leavesArea(228394000, Area1166285600), 20)
HappensAt(velocity_gt_2(228394000), 20)
...
```

Vertex: ? InitiatedAt(stopped(228394000), 50)

```
HappensAt(stop_start(228394000), 50)
HappensAt(slow_motion_end(228394000), 50)
HappensAt(isInArea(228394000, Area1166285600), 50)
HappensAt(velocity_lt_2(228394000), 50)
...
```

Graph Construction

Vertex: InitiatedAt(stopped(228394000), 5)

```
HappensAt(stop_start(228394000), 5)
HappensAt(slow_motion_end(228394000), 5)
HappensAt(isInArea(228394000, Area1166285600), 5)
HappensAt(velocity_lt_2(228394000), 5)
...
```

Vertex: \neg InitiatedAt(stopped(228394000), 20)

```
HappensAt(stop_end(228394000), 20)
HappensAt(slow_motion_start(228394000), 20)
HappensAt(leavesArea(228394000, Area1166285600), 20)
HappensAt(velocity_gt_2(228394000), 20)
...
```

Vertex: ? InitiatedAt(stopped(228394000), 50)

```
HappensAt(stop_start(228394000), 50)
HappensAt(slow_motion_end(228394000), 50)
HappensAt(isInArea(228394000, Area1166285600), 50)
HappensAt(velocity_lt_2(228394000), 50)
...
```

Graph Construction



Graph Construction



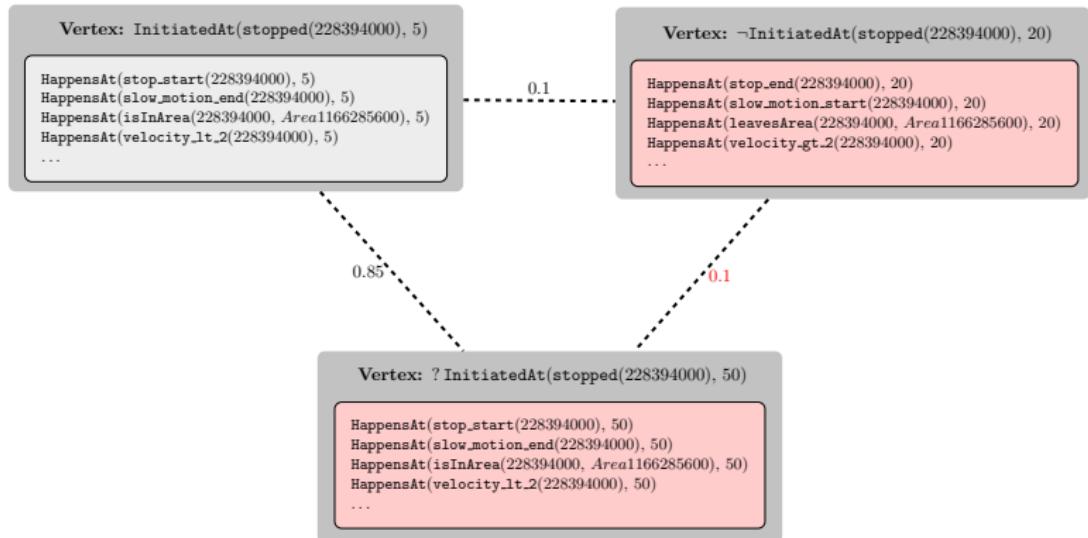
Graph Construction



Graph Construction



Graph Construction



Supervision Completion



Supervision Completion

Vertex: InitiatedAt(stopped(228394000), 5)

```
HappensAt(stop_start(228394000), 5)
HappensAt(slow_motion_end(228394000), 5)
HappensAt(isInArea(228394000, Area1166285600), 5)
HappensAt(velocity_lt_2(228394000), 5)
...
```

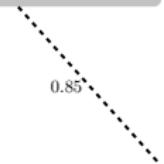
Vertex: \neg InitiatedAt(stopped(228394000), 20)

```
HappensAt(stop_end(228394000), 20)
HappensAt(slow_motion_start(228394000), 20)
HappensAt(leavesArea(228394000, Area1166285600), 20)
HappensAt(velocity_gt_2(228394000), 20)
...
```

Vertex: ? InitiatedAt(stopped(228394000), 50)

```
HappensAt(stop_start(228394000), 50)
HappensAt(slow_motion_end(228394000), 50)
HappensAt(isInArea(228394000, Area1166285600), 50)
HappensAt(velocity_lt_2(228394000), 50)
...
```

0.85



Supervision Completion

Vertex: InitiatedAt(stopped(228394000), 5)

```
HappensAt(stop_start(228394000), 5)
HappensAt(slow_motion_end(228394000), 5)
HappensAt(isInArea(228394000, Area1166285600), 5)
HappensAt(velocity_lt_2(228394000), 5)
...
```

Vertex: \neg InitiatedAt(stopped(228394000), 20)

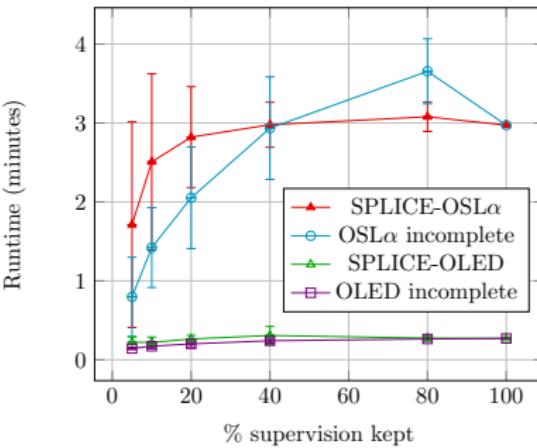
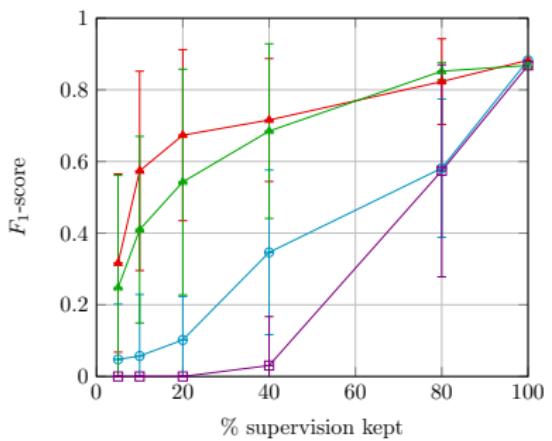
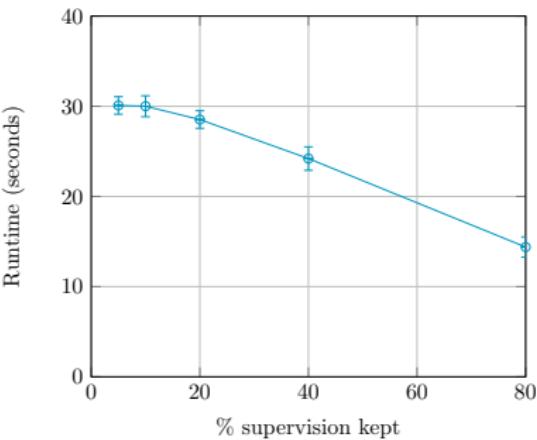
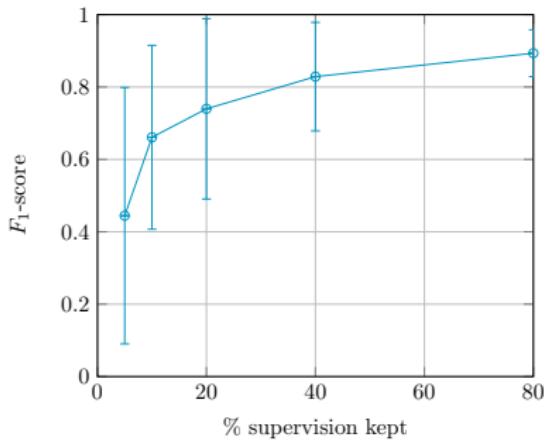
```
HappensAt(stop_end(228394000), 20)
HappensAt(slow_motion_start(228394000), 20)
HappensAt(leavesArea(228394000, Area1166285600), 20)
HappensAt(velocity_gt_2(228394000), 20)
...
```

0.85

Vertex: InitiatedAt(stopped(228394000), 50)

```
HappensAt(stop_start(228394000), 50)
HappensAt(slow_motion_end(228394000), 50)
HappensAt(isInArea(228394000, Area1166285600), 50)
HappensAt(velocity_lt_2(228394000), 50)
...
```

Experimental Evaluation (1/2)



Experimental Evaluation (2/2)

