

# Human-Centered ML

## Part II



**HUMAN-CENTERED MACHINE LEARNING**

<http://courses.mpi-sws.org/hcml-ws18/>

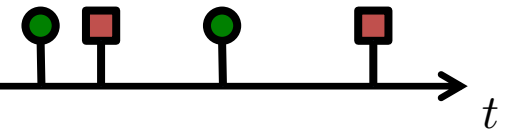


MAX PLANCK INSTITUTE  
FOR SOFTWARE SYSTEMS

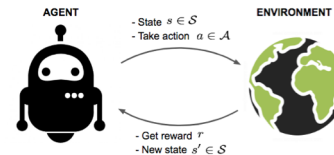
# Lectures for Part II

## Five lectures on fundamentals

→ Marked temporal Point Processes



→ Optimal control  
Reinforcement learning



## Five lectures on applications



Information  
propagation



Viral  
marketing



Opinion  
mining



Information  
integrity



Human  
learning

# Evaluation for Part II

## **Paper reviewing assignments:**

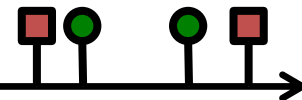
- **Only for lectures on applications**
- **Due just before the lecture**

## **Two coding assignments:**

- **Information propagation**  
From Dec 20 to Jan 17
- **Viral marketing**  
From Jan 24 to Feb 5

**Final exam:** On Feb 7 (review on Feb 5)

# Introduction to Temporal Point Processes (I)



HUMAN-CENTERED MACHINE LEARNING

<http://courses.mpi-sws.org/hcml-ws18/>



MAX PLANCK INSTITUTE  
FOR SOFTWARE SYSTEMS

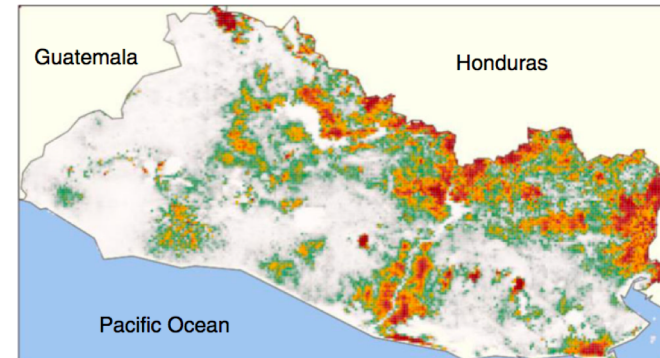
# Many discrete *events* in continuous time



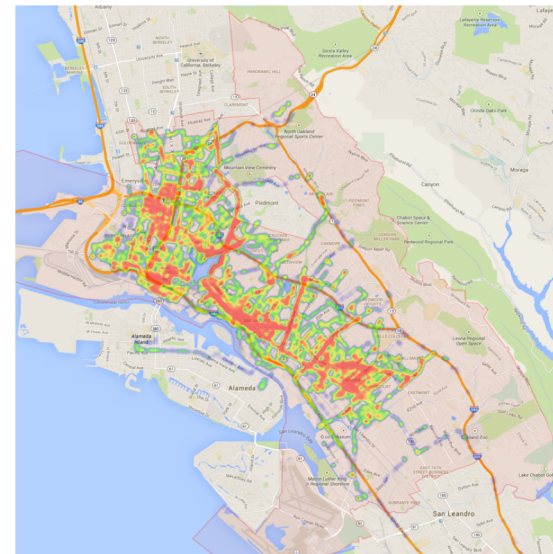
**Online actions**



**Financial trading**



**Disease dynamics**



**Mobility dynamics**

# Variety of processes behind these events

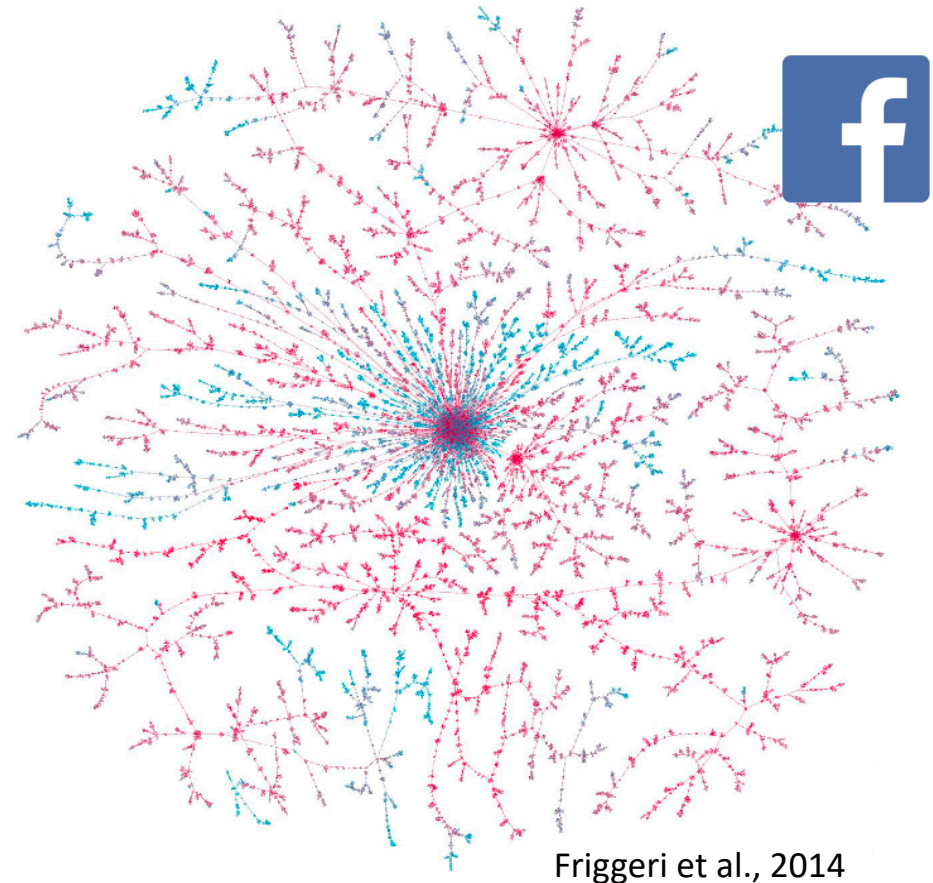
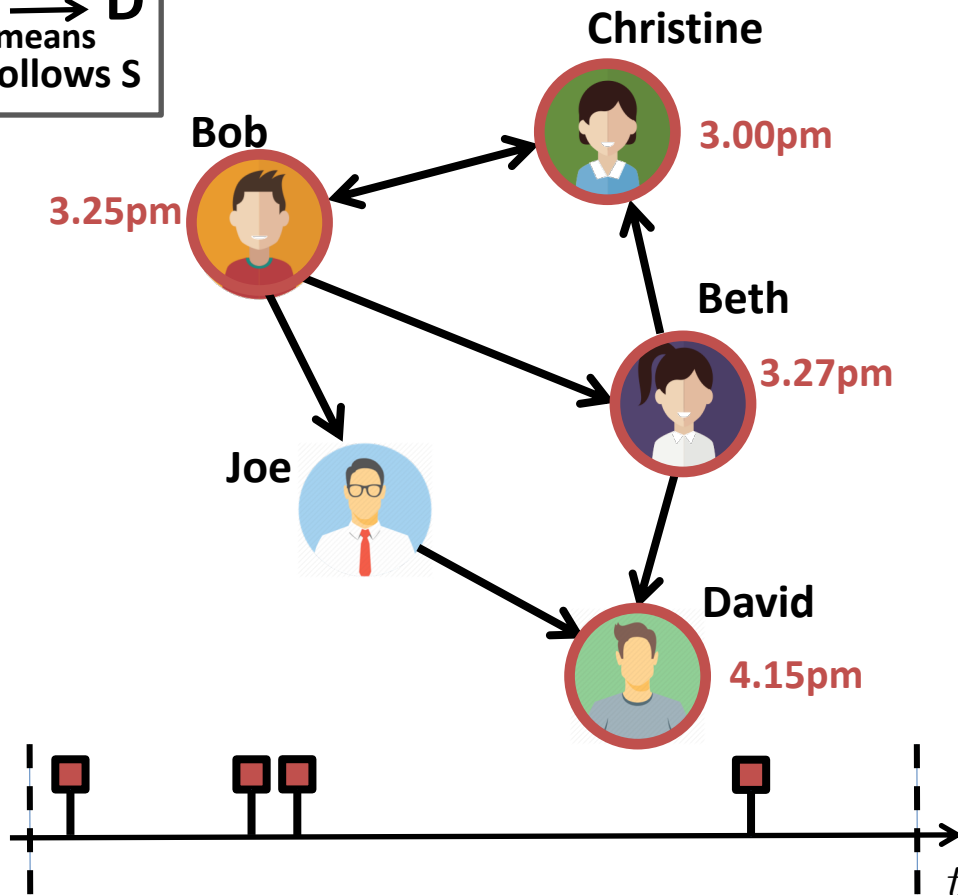
Events are (noisy) observations of a variety of complex dynamic processes...



...in a wide range of temporal scales. 6

# Example I: Information propagation

$S \rightarrow D$   
means  
D follows S



They can have an impact  
in the off-line world

**theguardian**

Click and elect: how fake news helped Donald Trump win a real election

# Example II: Knowledge creation



Barack Obama

From Wikipedia, the free encyclopedia

*"Barack" and "Obama" redirect here. For his father, see Barack Obama Sr. For other uses of "Barack", see Barack (disambiguation). (disambiguation).*

**Barack Hussein Obama II** (), current President of the United States. He was president of the *Harvard* civil rights attorney and taught representing the 13th District States House of Representatives

## Barack Obama: Revision history

03:41, 28 November 2016	Ranze (talk   contribs)	.. (301,105 bytes) (+18) .. (E
03:32, 28 November 2016	Xin Deui (talk   contribs)	.. (301,087 bytes) (-68) .. (
00:57, 28 November 2016	SporkBot (talk   contribs)	m .. (301,155 bytes) (-37)
07:03, 27 November 2016	Saiph121 (talk   contribs)	.. (301,192 bytes) (+25) ..

03:21, 20 September 2016

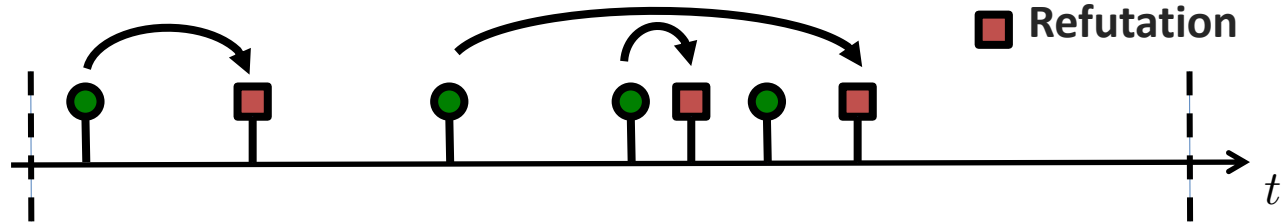
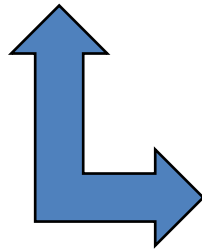
is a Kenyan politician



*possible vandalism by MLM2016*

is an American politician

● Addition  
■ Refutation



Moving to Australia Working in Australia Study abroad in Australia +4

## What are the pros and cons of living in Australia?

Answer Request Follow 109 Comment Share 9 Downvote

*I have studied, worked and lived in Australia as an Internat...  
employee, business owner and a citizen.*

Upvote | 150

*I have experienced this country in all the ways possible, you  
However, I firmly believe that there are definitely more pros  
Australia but still I have mentioned below a few...*



...v M Sharma, Lived in Australia as Migrant, Student, Worker,  
Business Owner & Family Man

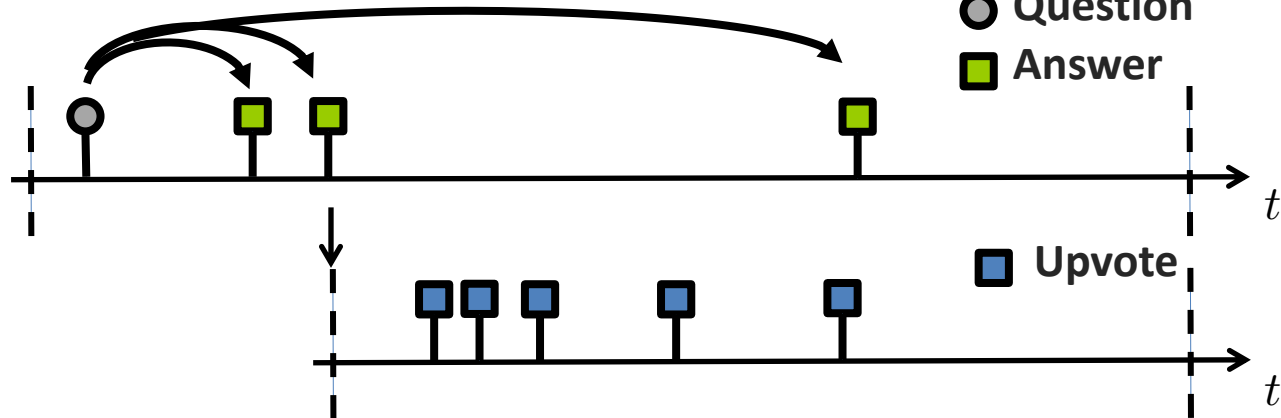
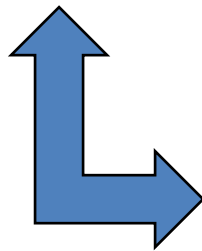
Updated Aug 3

Hope it helped :)

Possible Challenges

- Language problem for those who don't speak...
- Not having your family and friends around coul...
- society is more and more connected and thanks...
- Social Media you can stay in touch a bit easier w...

● Question  
■ Answer



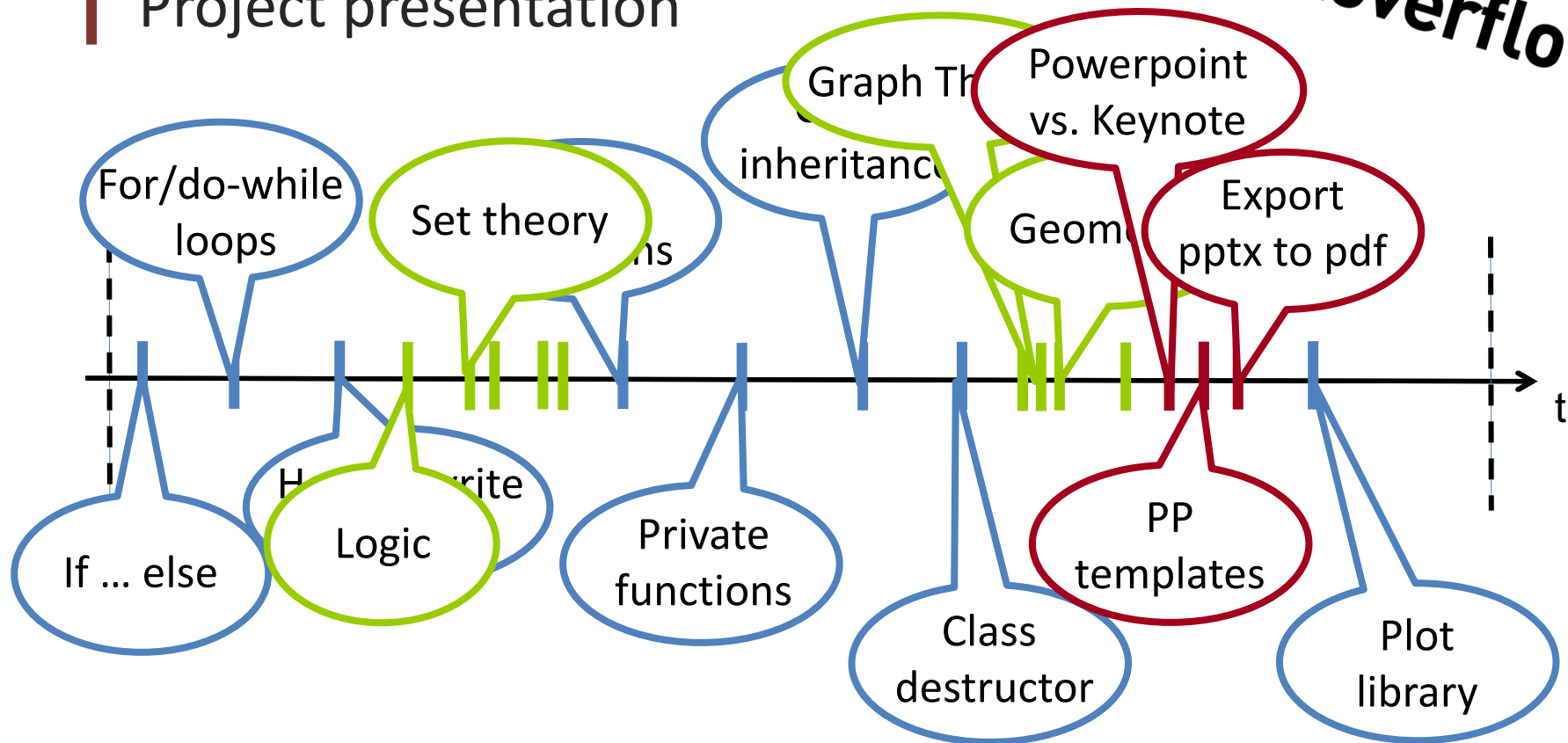


# Example III: Human learning



## 1st year computer science student

- Introduction to programming
- Discrete math
- Project presentation



# Detailed *event traces*

DETAILED TRACES OF ACTIVITY



**Warren Buffett** ✓  
@WarrenBuffett



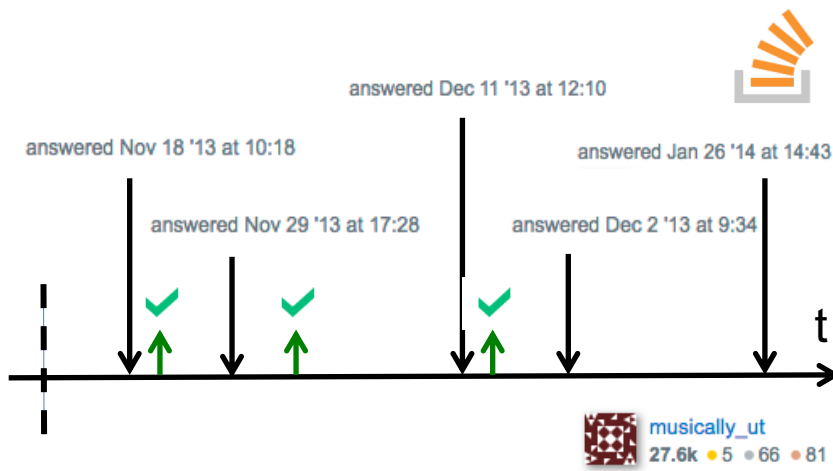
Warren is in the house.

↩ Reply ↻ Retweet ★ Favorite ⋮ More

RETWEETS 43,649 FAVORITES 14,328



9:20 AM - 2 May 2013



**Manuel Gomez Rodriguez** updated his cover photo.

April 17 at 1:14pm · ☺



Pique-Longue, French Pyrenees  
Easter 2017



👍 Like 💬 Comment ➦ Share

👍❤️ Mehrdad Farajtabar, Lili Yavis-Hound and 24 others



**Rober Tab Pu** 😊 wow!

Like · Reply · April 17 at 1:32pm

# Detailed *event traces*

DETAILED TRACES OF ACTIVITY



Warren Buffett ✓  
@WarrenBuffett



Warren is in the house.

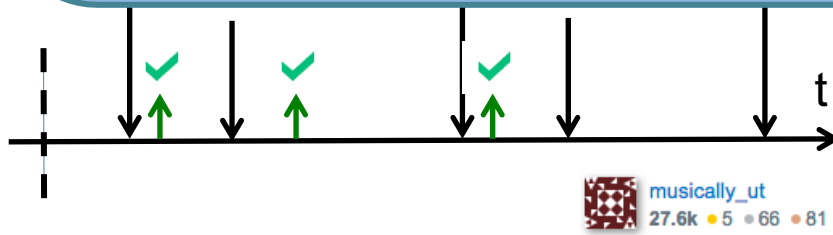


Manuel Gomez Rodriguez updated his cover photo.  
April 17 at 1:14pm · 🌐



Pique-Longue, French Pyrenees  
Easter 2017

The availability of event traces  
boosts a new generation of  
data-driven models and algorithms



Like Comment Share

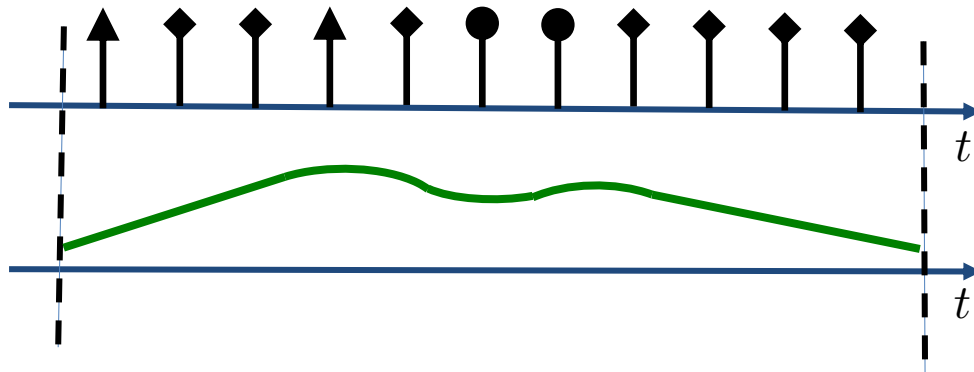
Mehrdad Farajtabar, Lili Yavis-Hound and 24 others



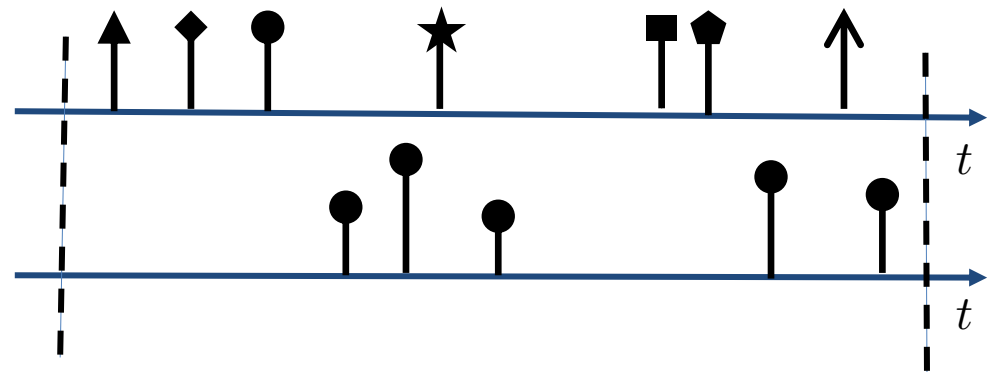
Rober Tab Pu 😲 wow!

Like · Reply · April 17 at 1:32pm

# Aren't these event traces just time series?

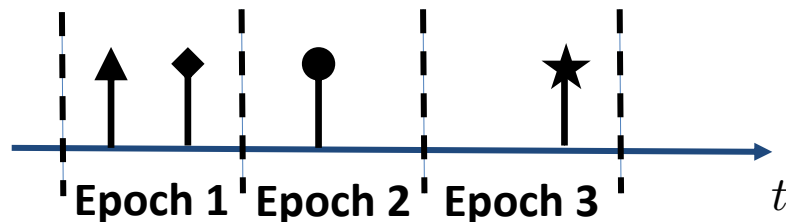


**Discrete and continuous times series**



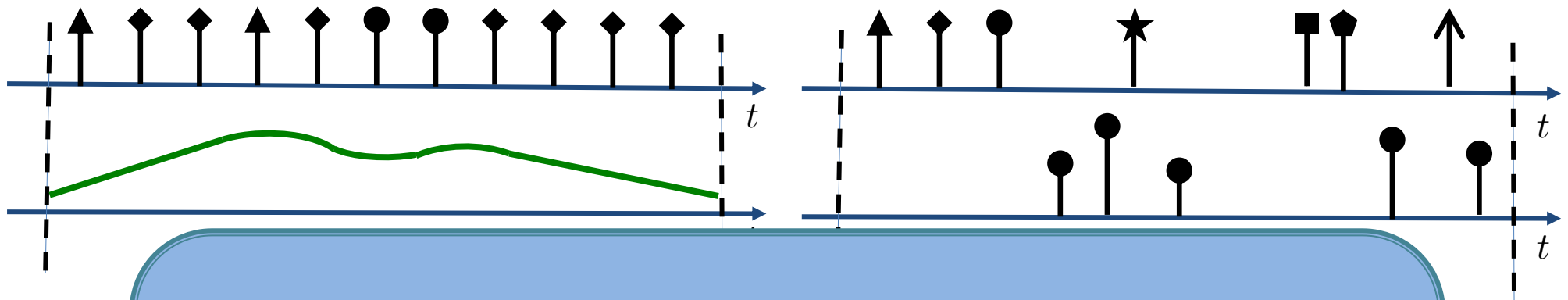
**Discrete events in continuous time**

**What about aggregating events in *epochs*?**



- How long is each epoch?
- How to aggregate events per epoch?
- What if no event in one epoch?
- What about time-related queries?

# Aren't these event traces just time series?



Dis

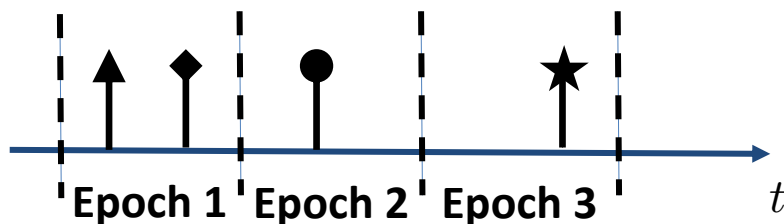
W

The framework of  
**temporal point processes**  
provides a *native representation*

epoch?

What if no event in one epoch?

What about time-related queries?

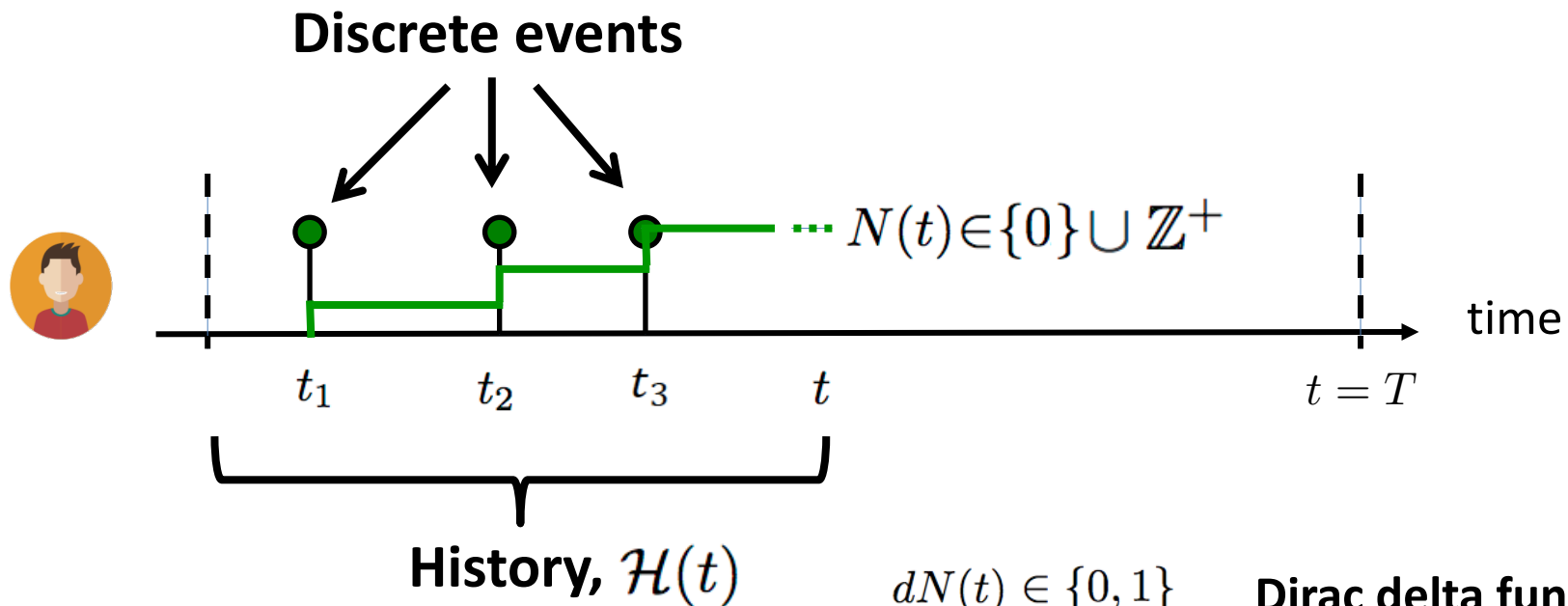


# Temporal Point Processes: Intensity function

# Temporal point processes

## Temporal point process:

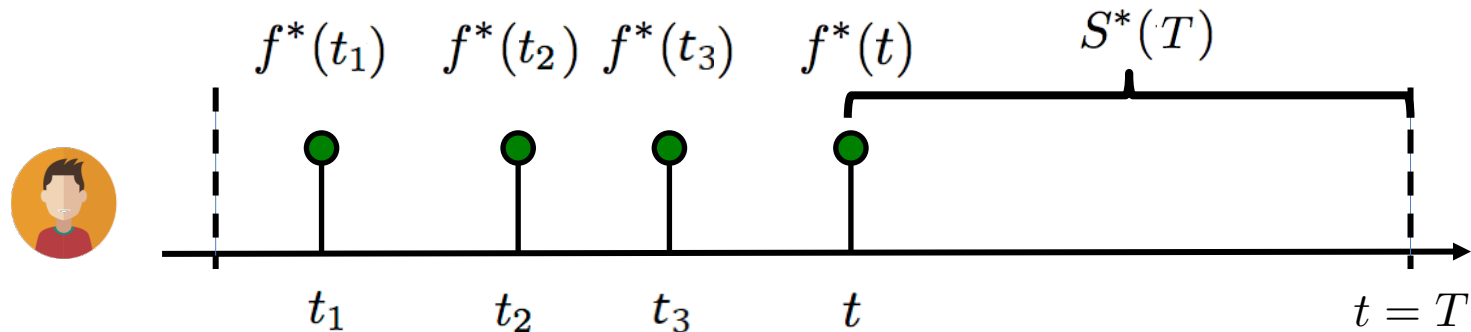
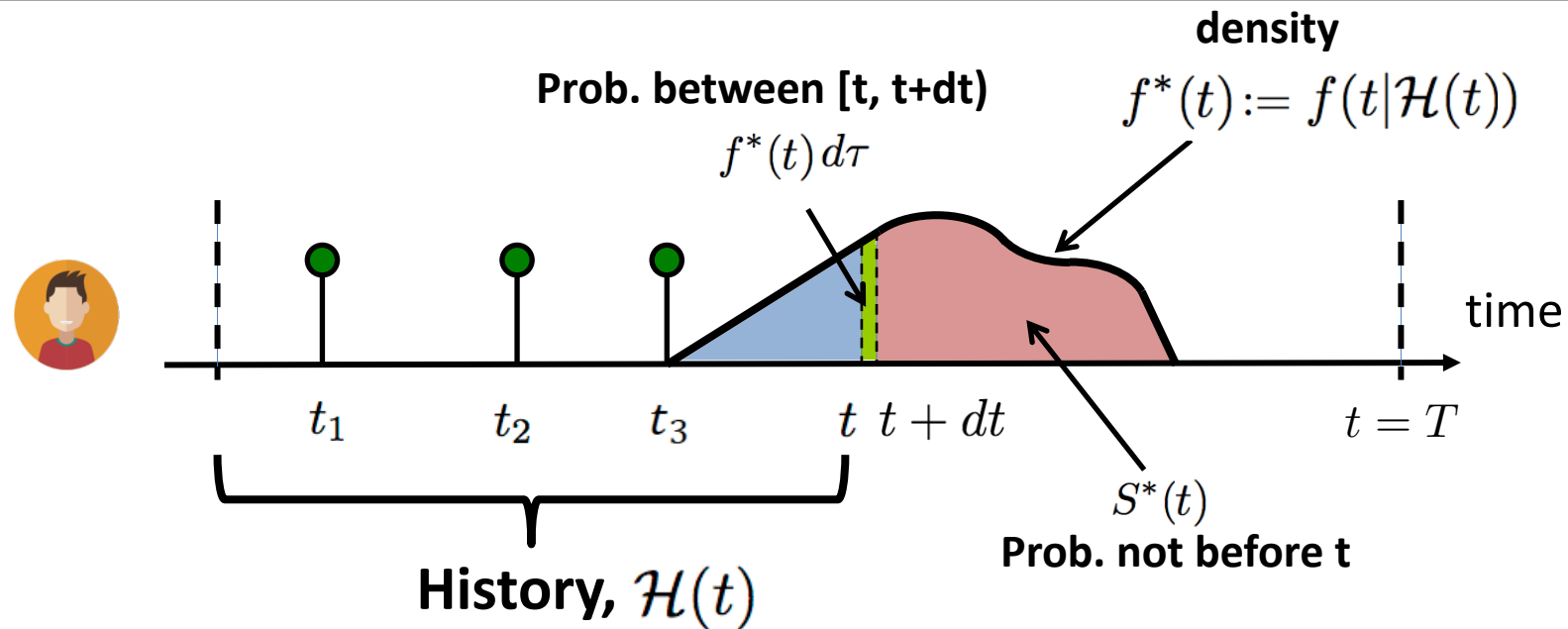
A random process whose realization consists of discrete events localized in time



**Formally:**  $N(t) = \int_0^t dN(s) \Rightarrow dN(t) = \sum_{t_i \in \mathcal{H}(t)} \delta(t - t_i) dt$

$dN(t) \in \{0, 1\}$       Dirac delta function

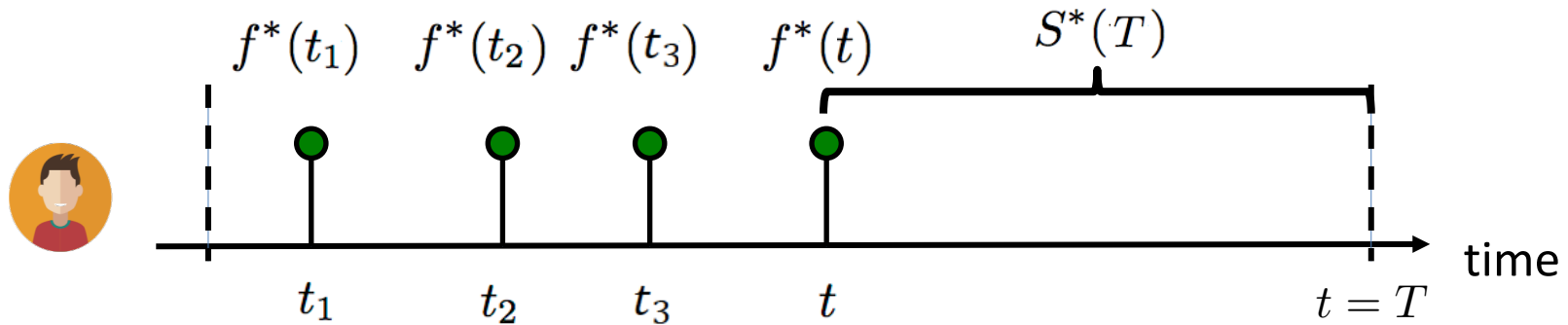
# Model time as a random variable



Likelihood of a timeline:  $f^*(t_1) f^*(t_2) f^*(t_3) f^*(t) S^*(T)$



# Problems of density parametrization (I)



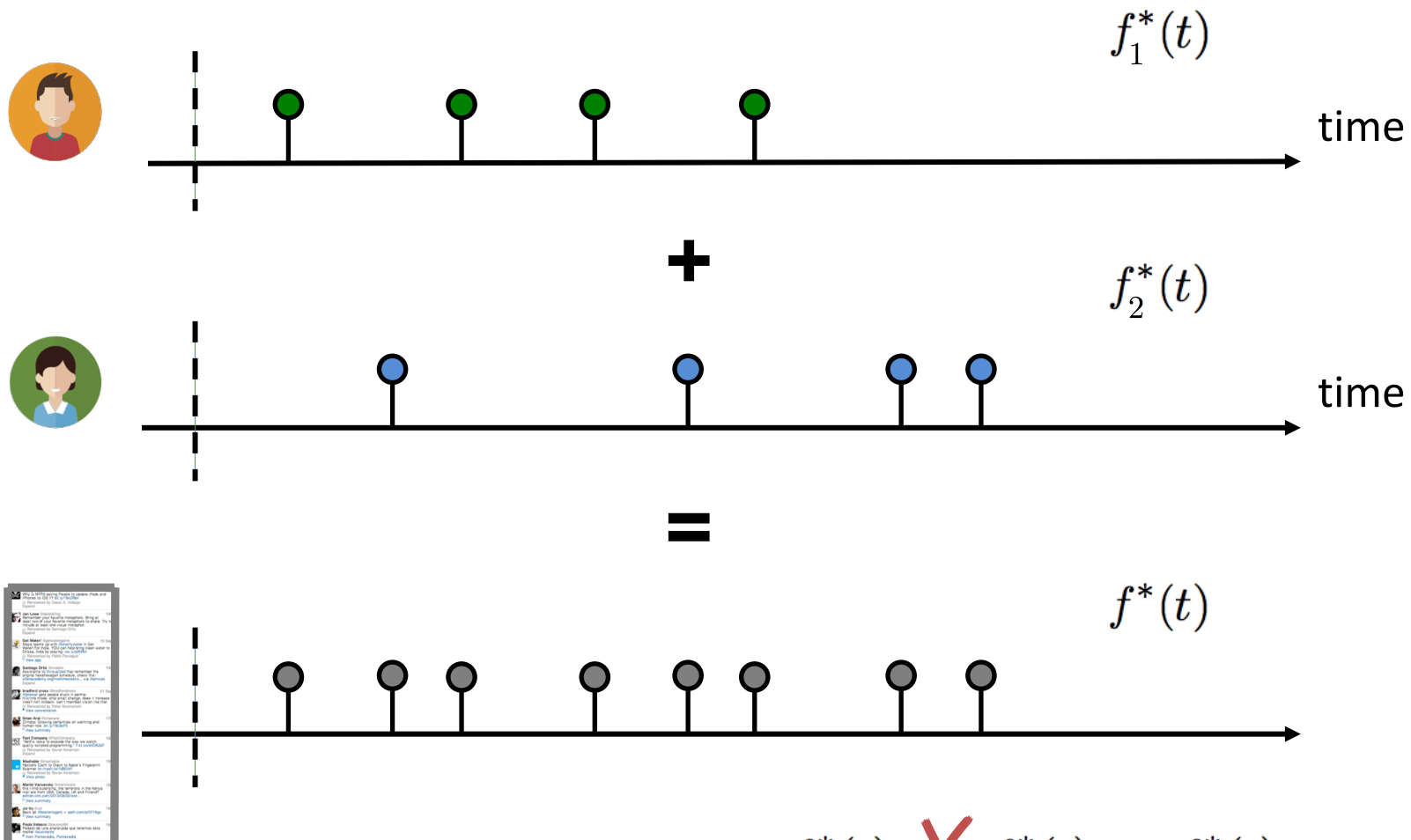
$$\begin{array}{ccccccc}
 f^*(t_1) & f^*(t_2) & f^*(t_3) & f^*(t) & S^*(T) & & \\
 \nearrow & \nearrow & \uparrow & \nwarrow & \nwarrow & & \\
 \frac{\exp\langle w, \psi^*(t_1) \rangle}{Z} & & \frac{\exp\langle w, \psi^*(t_3) \rangle}{Z} & & & & 1 - \int_t^T \frac{\exp\langle w, \psi^*(\tau) \rangle}{Z} d\tau \\
 & \frac{\exp\langle w, \psi^*(t_2) \rangle}{Z} & & \frac{\exp\langle w, \psi^*(t) \rangle}{Z} & & & 
 \end{array}$$

It is **difficult for model design and interpretability**:

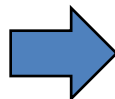
1. Densities need to integrate to 1 (i.e., partition function)
2. Difficult to combine timelines

# Problems of density parametrization (II)

Difficult to combine timelines:

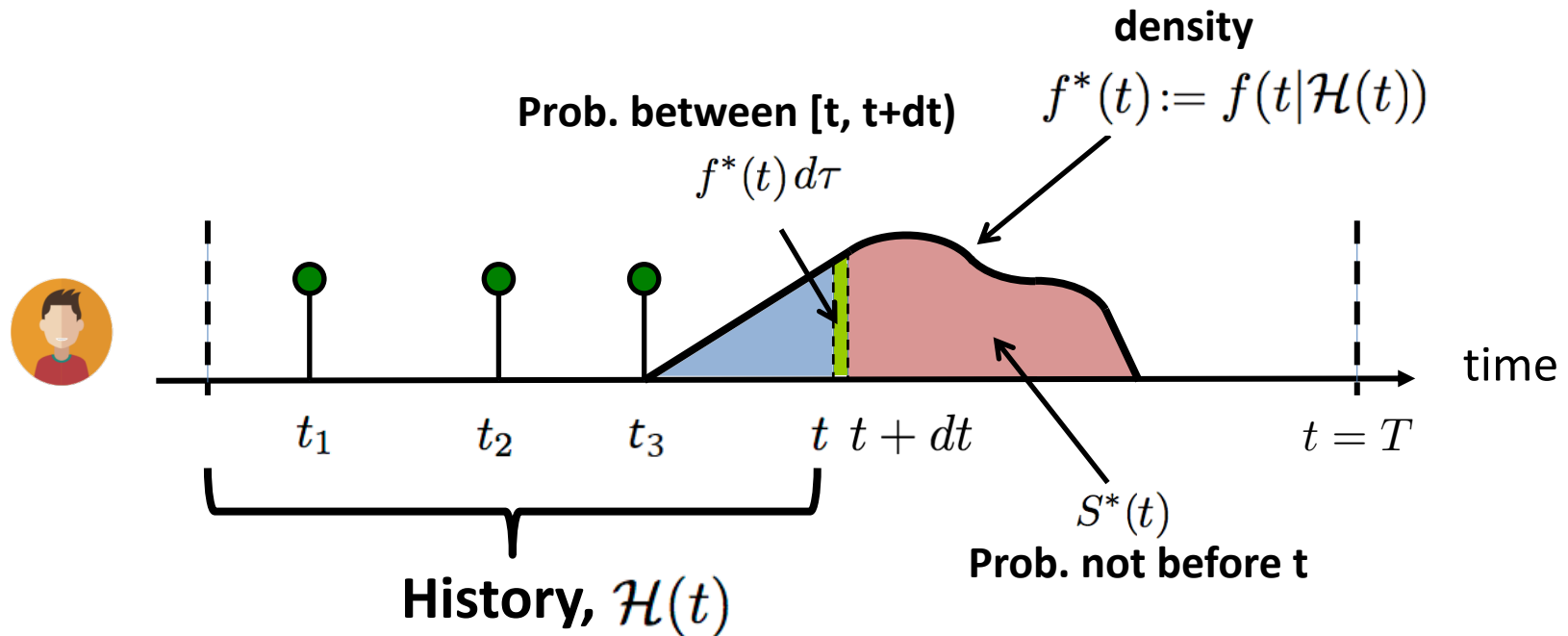


Sum of random processes



$$f^*(t) \neq f_1^*(t) + f_2^*(t)$$
$$f^*(t) \neq f_1^*(t) * f_2^*(t)$$

# Intensity function



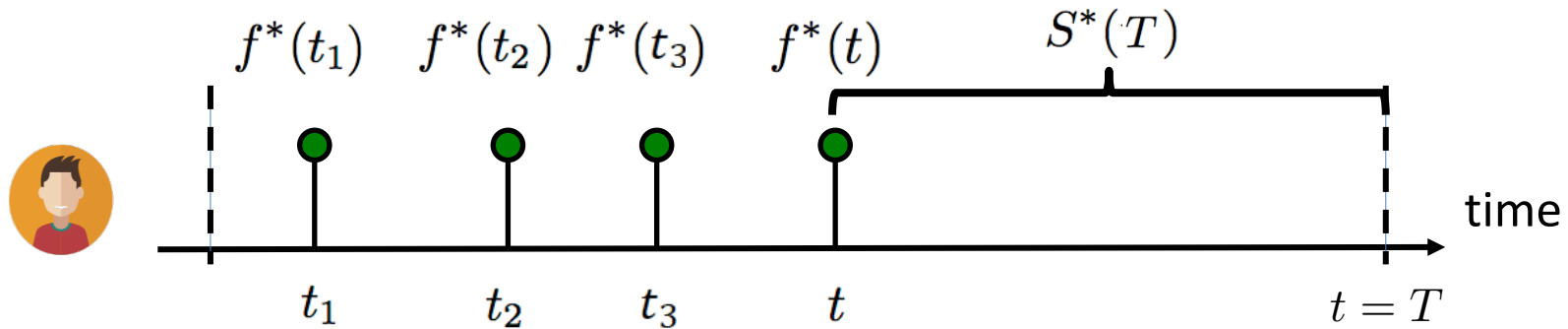
**Intensity:**

Probability between  $[t, t+dt)$  but not before t

$$\lambda^*(t)dt = \frac{f^*(t)dt}{S^*(t)} \geq 0 \quad \Rightarrow \quad \lambda^*(t)dt = \mathbb{E}[dN(t)|\mathcal{H}(t)]$$

**Observation:**  $\lambda^*(t)$  It is a rate = # of events / unit of time

# Advantages of intensity parametrization (I)



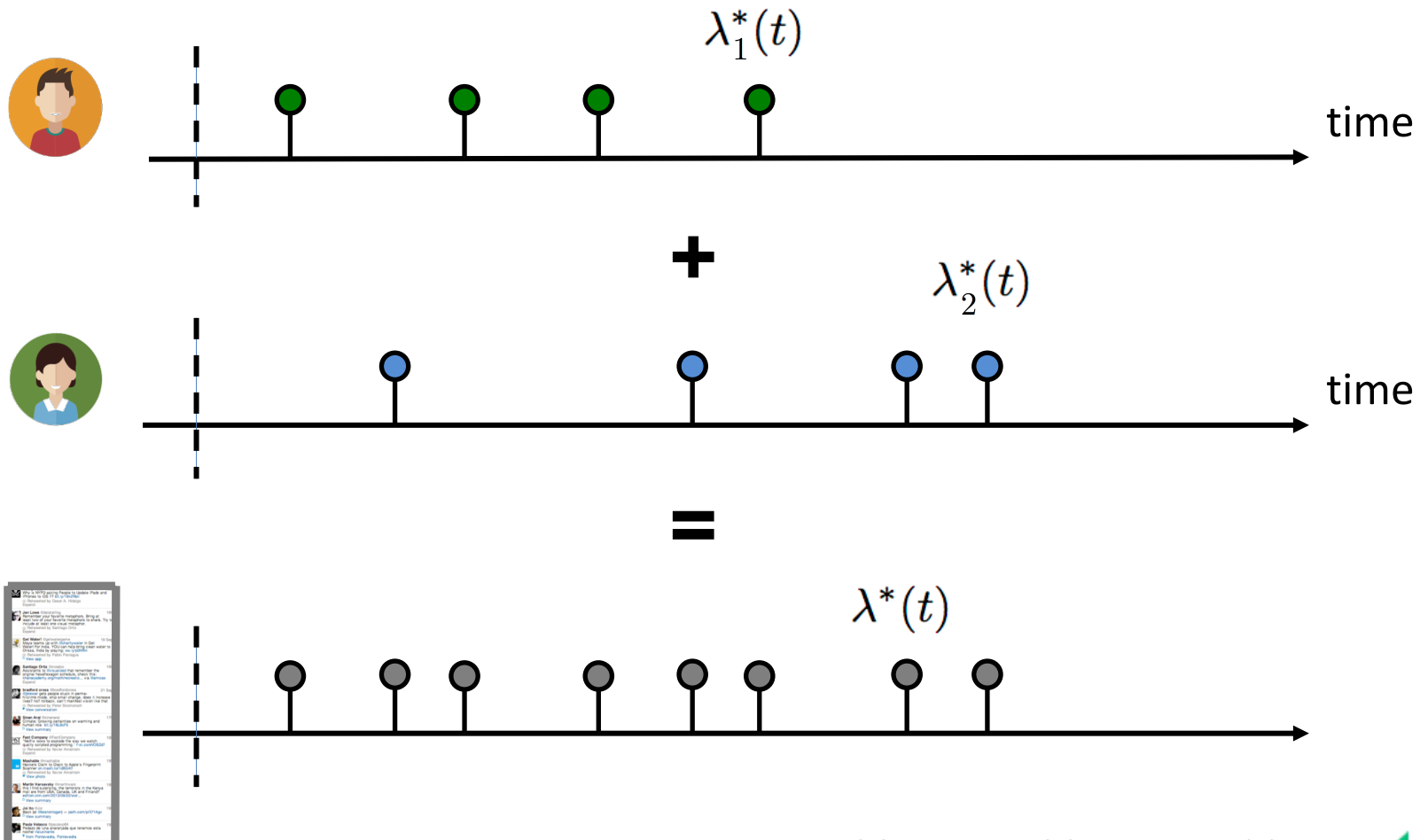
$$\begin{array}{ccccccc}
 \lambda^*(t_1) & \lambda^*(t_2) & \lambda^*(t_3) & \lambda^*(t) & \exp\left(-\int_0^T \lambda^*(\tau) d\tau\right) & & \\
 \nearrow & \nearrow & \uparrow & \nwarrow & \nwarrow & & \\
 \langle w, \phi^*(t_1) \rangle & & \langle w, \phi^*(t_3) \rangle & & \langle w, \phi^*(t) \rangle & & \exp\left(-\int_0^T \langle w, \phi^*(\tau) \rangle d\tau\right) \\
 \nearrow & & & & & & \\
 \langle w, \phi^*(t_2) \rangle & & & & & & 
 \end{array}$$

**Suitable for model design and interpretable:**

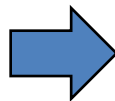
1. Intensities only need to be nonnegative
2. Easy to combine timelines

# Advantages of intensity parametrization (II)

Easy to combine timeline:



Sum of random processes



$$\lambda^*(t) = \lambda_1^*(t) + \lambda_2^*(t)$$



$$\lambda^*(t) \neq \lambda_1^*(t) * \lambda_2^*(t)$$

# Relation between $f^*$ , $F^*$ , $S^*$ , $\lambda^*$

