

# Applications of Conformal Prediction in Information Security Problems

Giovanni Cherubin

12 April 2016

# Machine Learning and Information Security

## Classification

- Attacks detection: malware, network traffic.
- Pre-analysis malware filtering.
- Attacks: traffic analysis, side channel.



IT geek



Student



Painter

# Machine Learning and Information Security

## Classification

- Attacks detection: malware, network traffic.
- Pre-analysis malware filtering.
- Attacks: traffic analysis, side channel.



IT geek



Student



Painter

## Anomaly detection

- Discover new threats.
- NIDS.



Normal



Normal



Anomaly

# Machine Learning and Information Security

## Classification

- Attacks detection: malware, network traffic.
- Pre-analysis malware filtering.
- Attacks: traffic analysis, side channel.



IT geek



Student



Painter

## Anomaly detection

- Discover new threats.
- NIDS.



Normal



Normal



Anomaly

## Clustering

- Monitoring tools.
- NIDS.



Cluster 1



Cluster 2

# Conformal Prediction

[VGS05]

- A statistical framework to make predictions.
- Controls the number of error.
- Applications that require confidence.

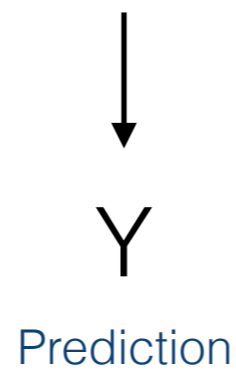
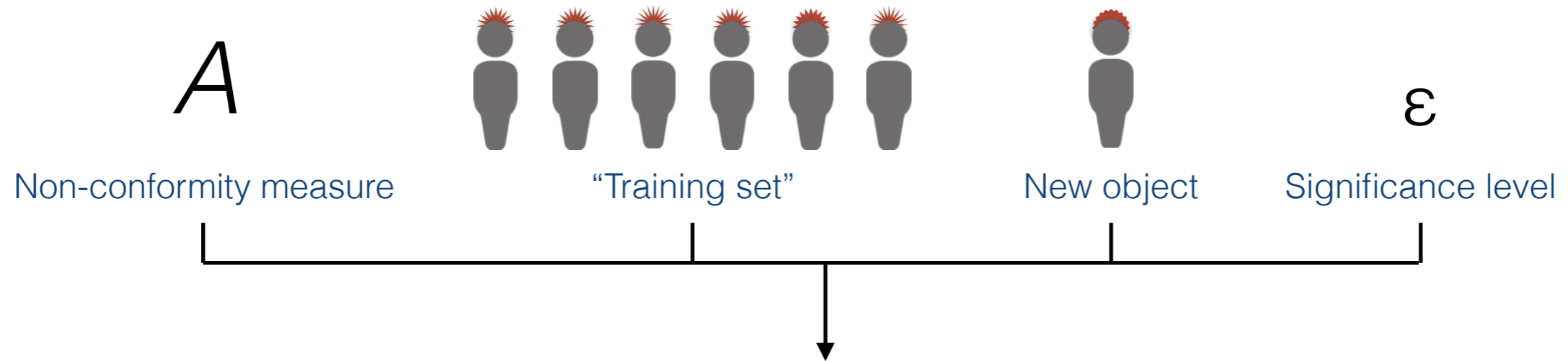
For **classification, anomaly detection, clustering, ...**

# Conformal Prediction

“non-conformity measure”

$$A(\text{[3 icons with large spikes]} ; \text{[1 icon with small spike]}) = 0.2$$

# Conformal Prediction



The error committed is smaller or equal to  $\epsilon$

# Classification of Malware

Conficker



Torpig



Zeus



New object



# Classification of Malware

Conficker



Torpig



Zeus



New object

# Classification of Malware

Conficker



Torpig



Zeus



New object

# Classification of Malware

Conficker



Torpig



Zeus



New object

# Classification of Malware

Conficker



Torpig



Zeus



New object

# Classification of Malware

Conficker



Torpig



Zeus



New object

$$Y = \{\text{Zeus}, \text{Torpig}\}$$

# Classification of Malware

Conficker



Torpig



Zeus



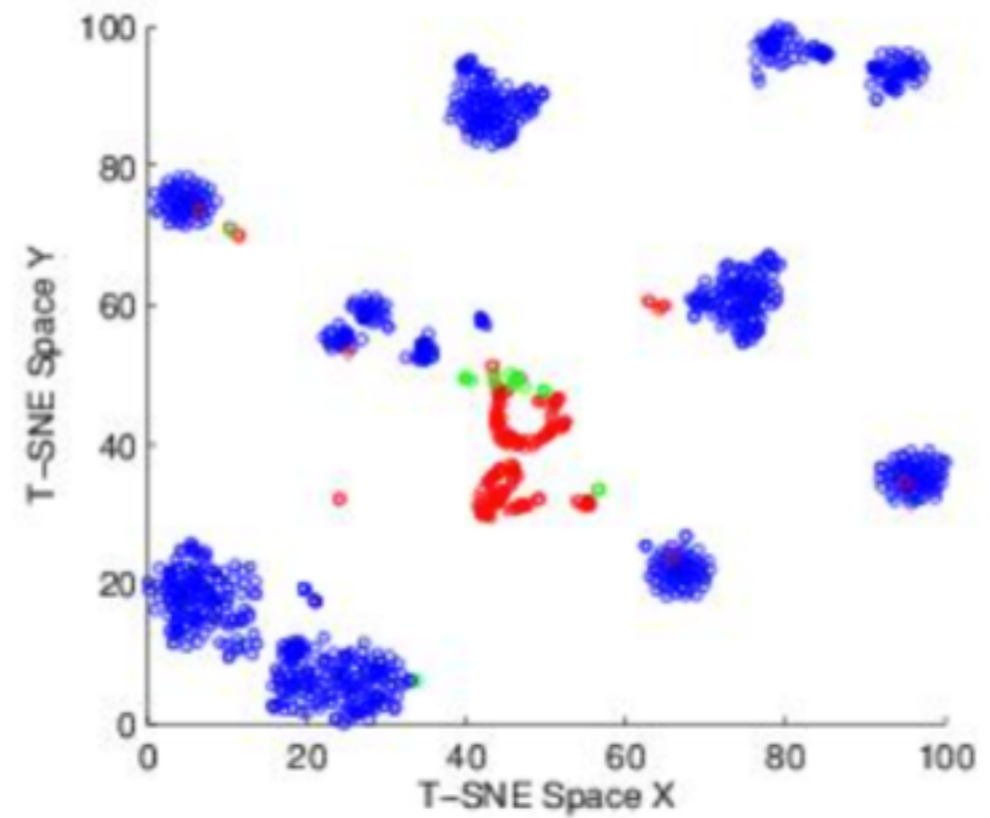
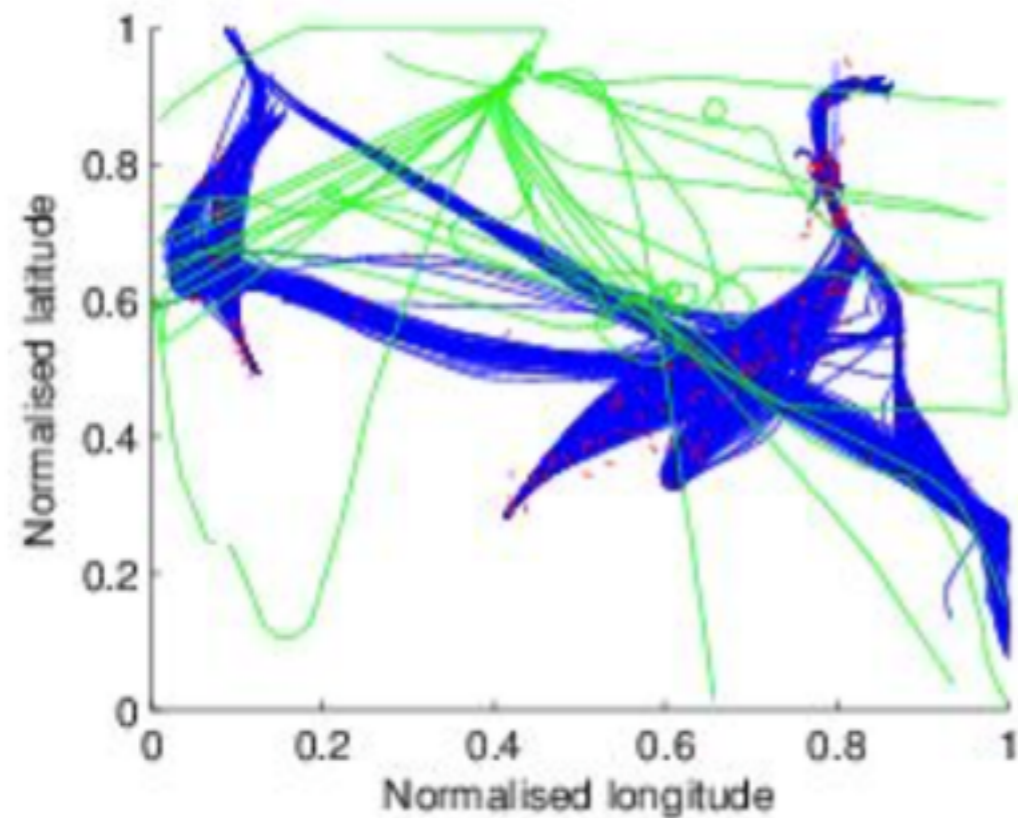
New object

$$Y = \{\text{Zeus}, \text{Torpig}\}$$

$$\text{Validity: } Pr(y \notin Y) \leq \epsilon$$

# Anomaly Detection

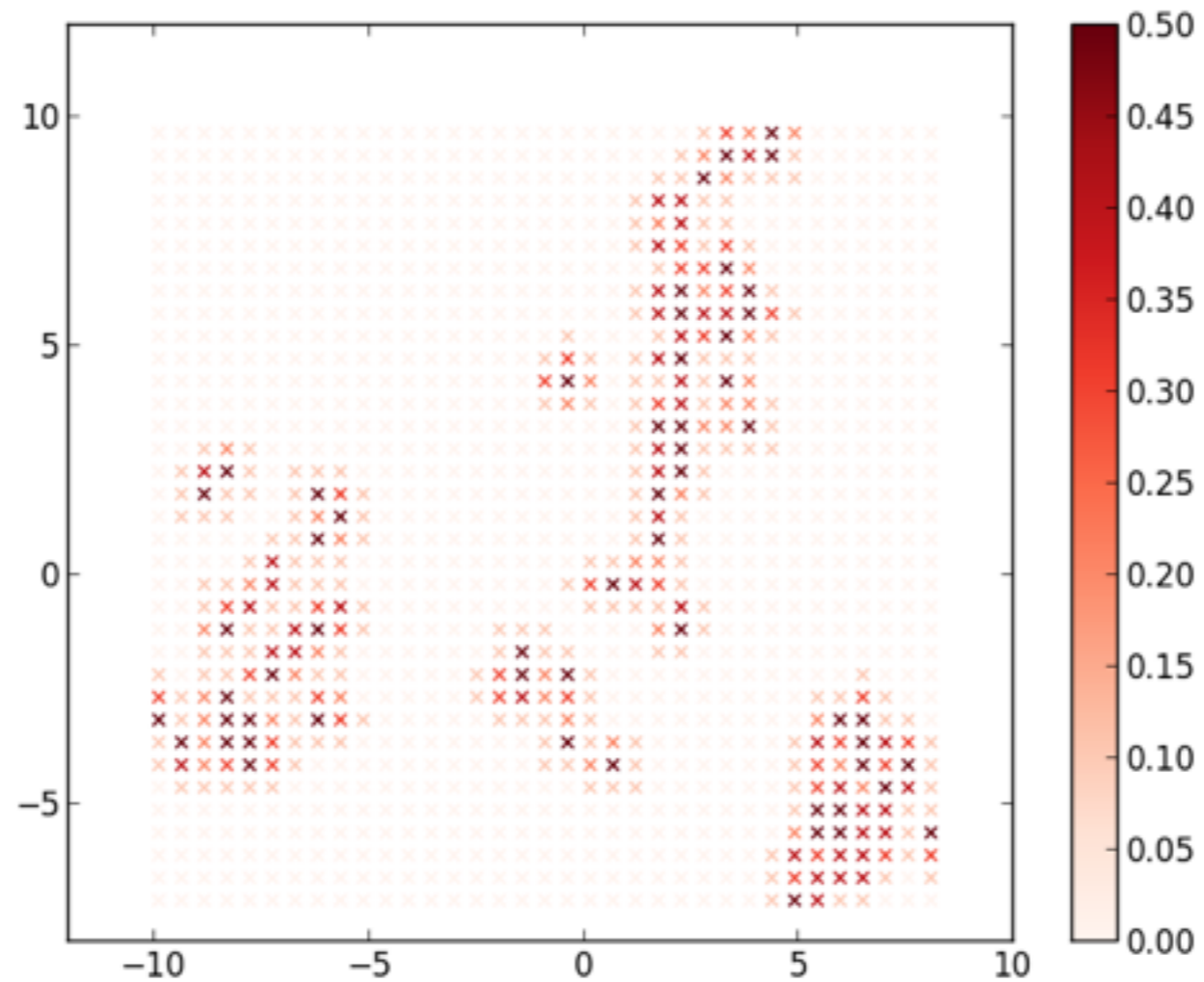
[SNCOG14]



Anomaly detection of maritime trajectories

# Bots clustering

[CNG+15]

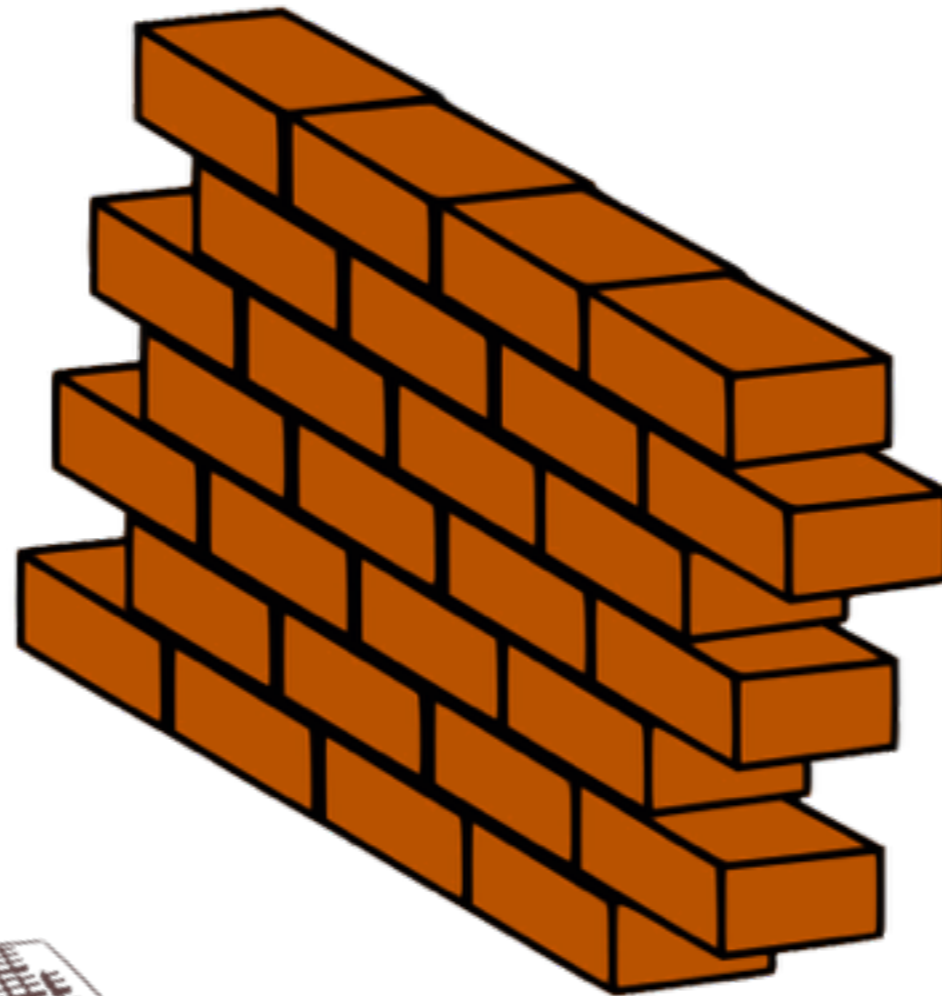


Forming clusters of similar bots network traces



# An attack...

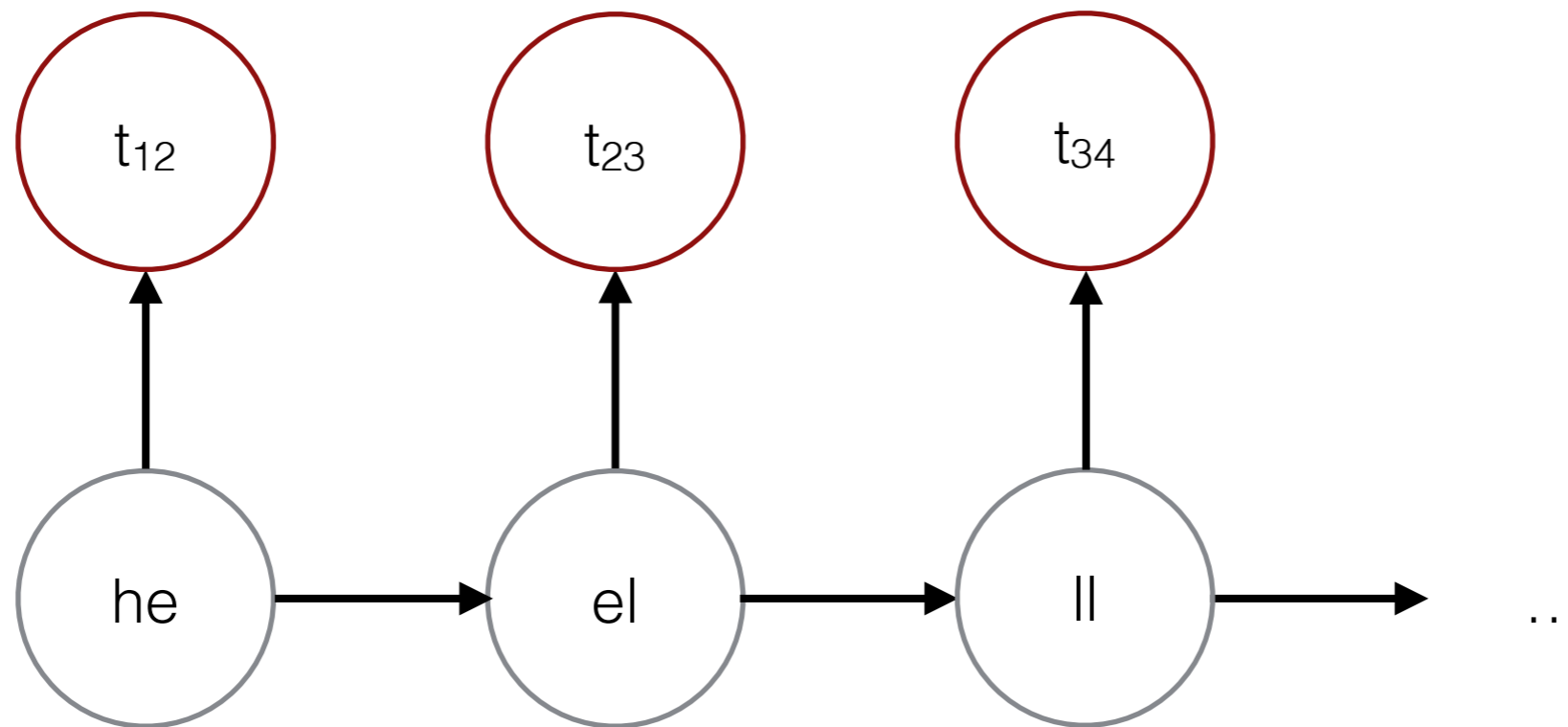
hello darling...



tic tic tic tic tic tic...

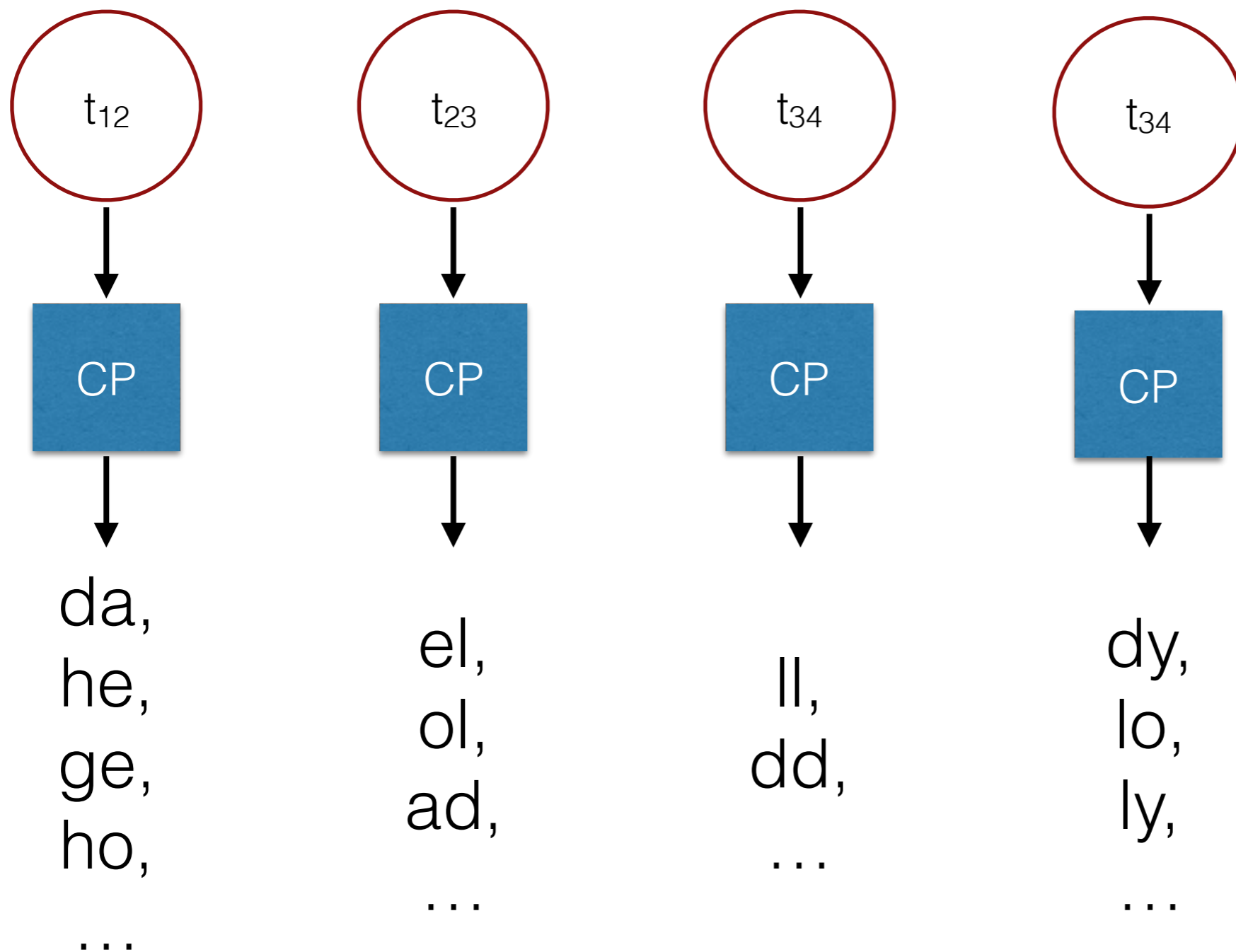
# Hidden Markov Models

tic .... tic



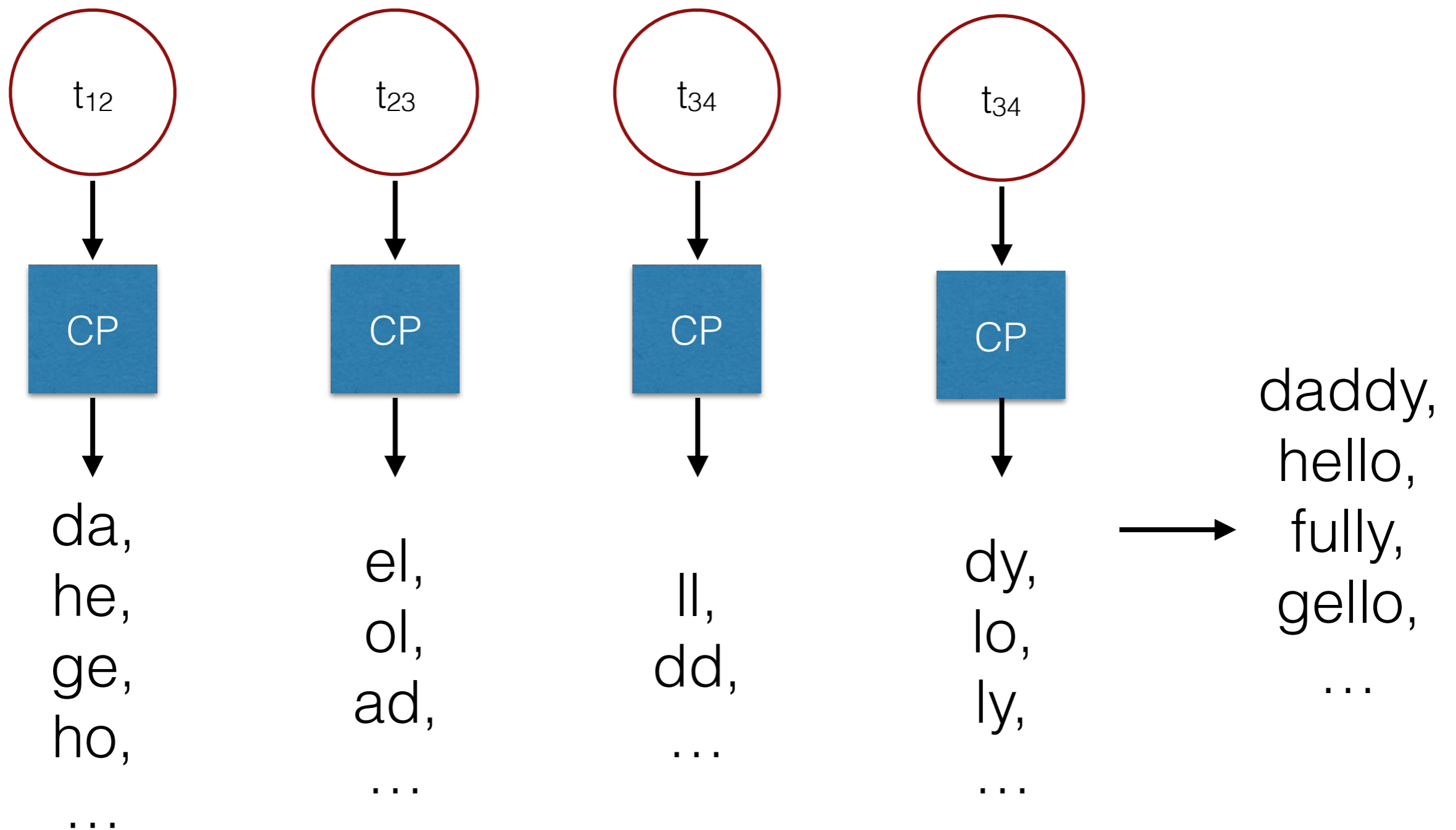
# Confident Prediction for HMM

[CN16]



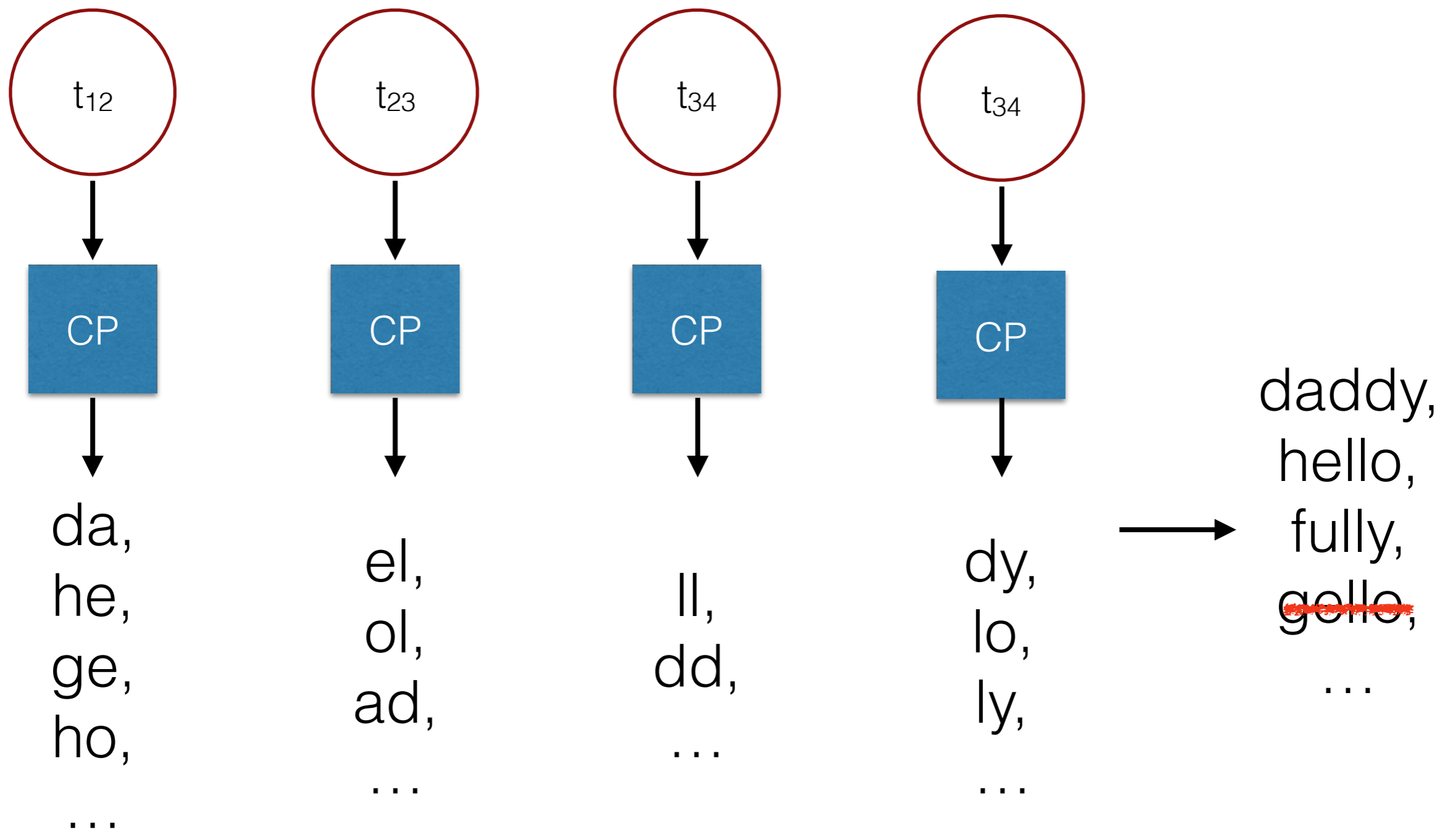
# Confident Prediction for HMM

[CN16]



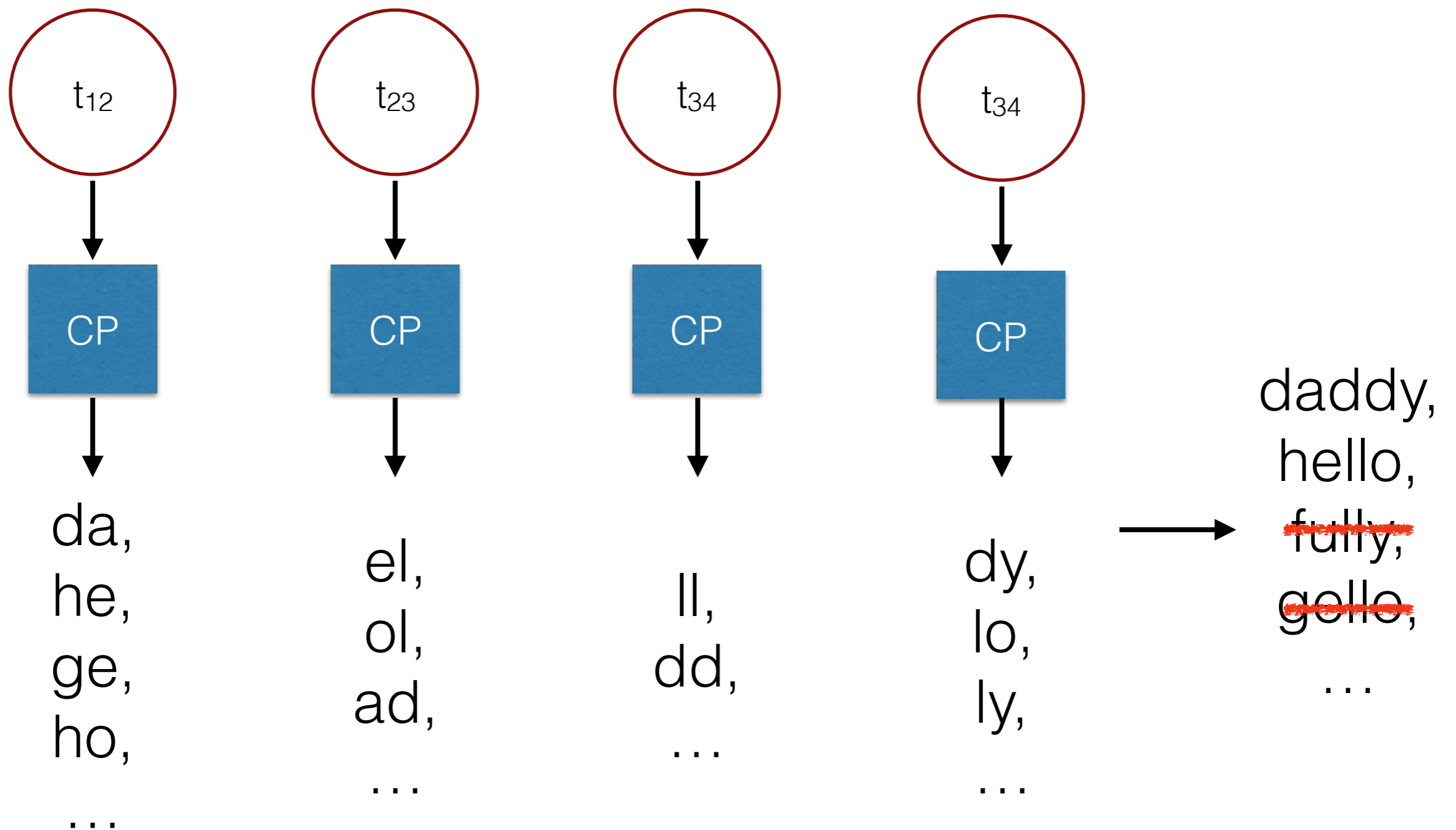
# Confident Prediction for HMM

[CN16]



# Confident Prediction for HMM

[CN16]



# Conclusions

- Conformal Prediction has interesting properties for Information security.
- It can be used for anomaly detection, classification, clustering, and beyond.
- Applications in Information Security exist, and much more can be done.

# References

- [VGS05] “Algorithmic learning in a random world”, V. Vovk, A. Gammernan, G. Shafer. 2005.
- [SNCOG14] “Anomaly Detection of Trajectories with Kernel Density Estimation by Conformal Prediction”, J. Smith, I. Nourtdinov, R. Craddock, C. Offer, A. Gammernan.
- [CNG<sup>+</sup>15] “Conformal Clustering and Its Application to Botnet Traffic”, G. Cherubin, I. Nourtdinov, A. Gammernan, R. Jordaney, Z. Wang, D. Papini, L. Cavallaro. 2015.
- [CN16] “Hidden Markov Models with Confidence”, G. Cherubin, I. Nourtdinov. 2016.

Conformal Prediction implementation: <https://github.com/gchers/cpy>.

Pictures from <https://openclipart.org>.



# Applications of Conformal Prediction in Information Security Problems

Giovanni Cherubin



@gchers

12 April 2016