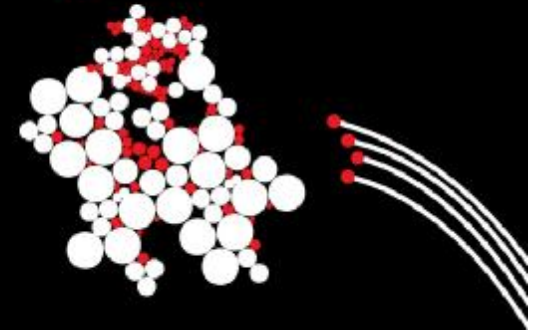
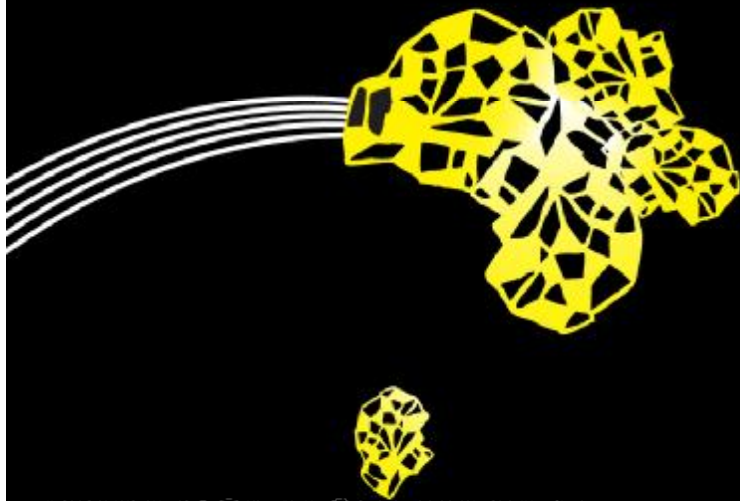


UNIVERSITEIT TWENTE.



# Handling Probabilistic Data

Ander de Keijzer



UNIVERSITY OF TWENTE.

# Medical domain

---

§ Physicians *are* inference systems

§ Physicians working together, means exchanging data

But....

§ Physicians really don't like uncertainty!!

## Brief overview of uncertain data

---

§ Data has an associated probability

Patient (name, diagnosis\*)

Amy, (Epilepsy || Parkinson)

## Brief overview of uncertain data

---

- § Data has an associated probability
- § Granularity of uncertainty depends on *level* with which the probability is associated

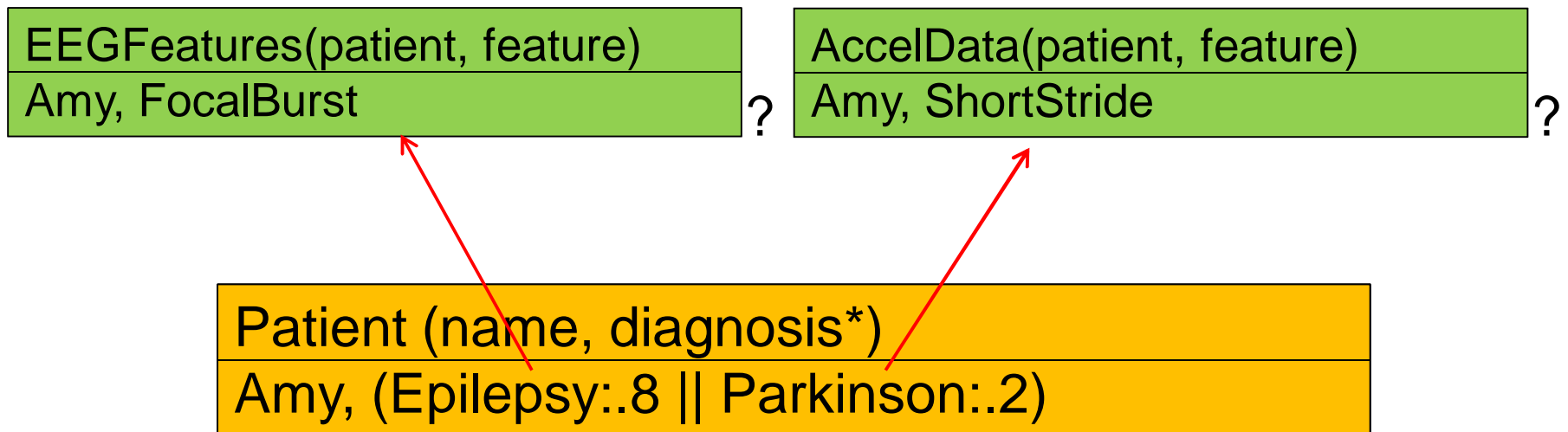
Patient (name, diagnosis\*)

Amy, (Epilepsy:.8 || Parkinson:.2)

## Brief overview of uncertain data

---

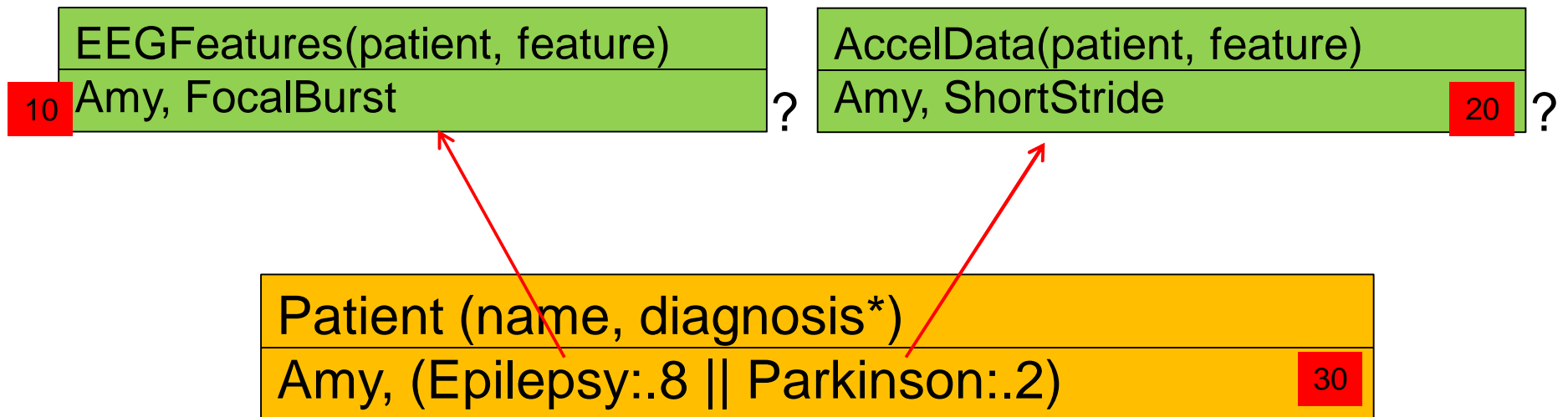
- § Data has an associated probability
- § Granularity of uncertainty depends on *level* with which the probability is associated



## Brief overview of uncertain data

---

- § Data has an associated probability
- § Granularity of uncertainty depends on *level* with which the probability is associated

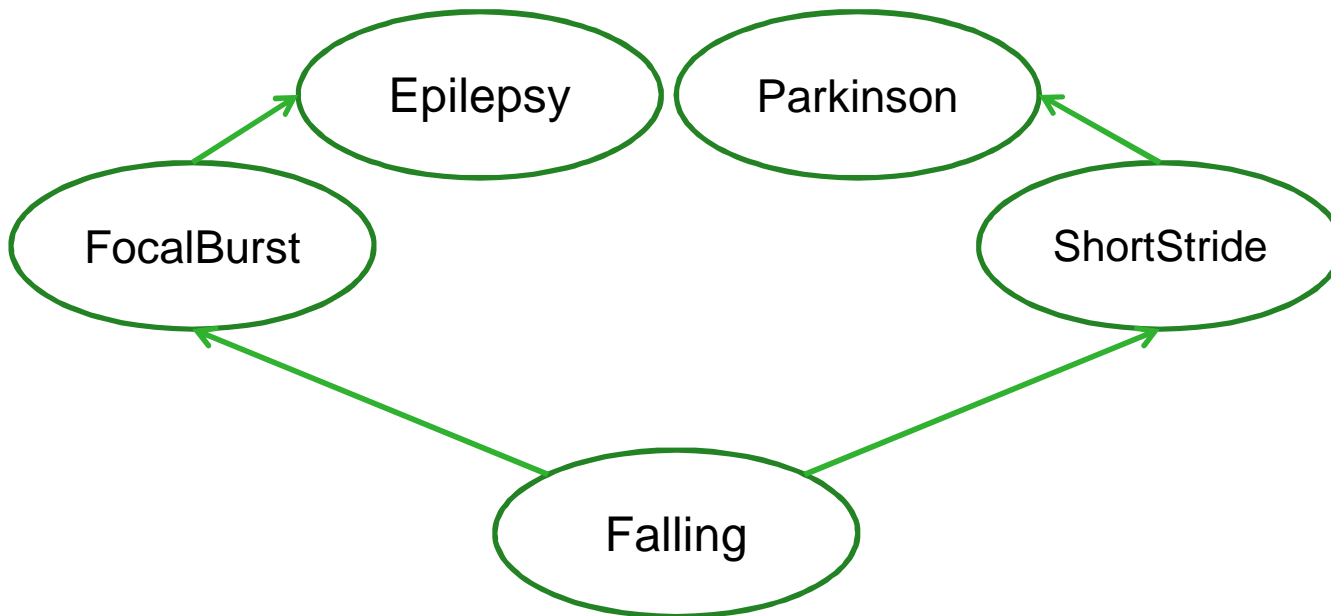


$$\lambda(30,1)=10; \lambda(30,2)=20$$

Or...

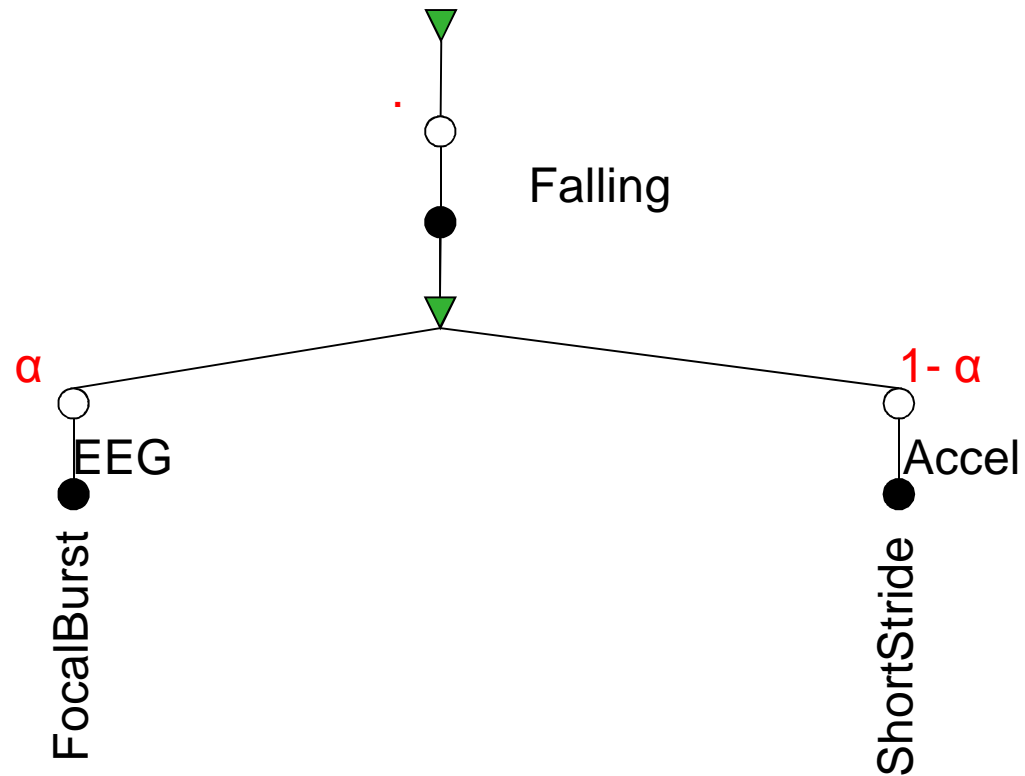
---

§ Probabilities given as conditional dependence probabilities, which incorporate lineage.



Or...

---





# Discussion

---

# Discussion

---

§ Where do the initial probabilities come from?

# Discussion

---

§ Where do the initial probabilities come from?

§ Are they even probabilities?? Or is it *confidence scores*?

# Discussion

---

- § Where do the initial probabilities come from?
  - § Are they even probabilities?? Or is it *confidence scores*?
- § Lineage...or provenance....or inference rules??
- § How to represent the data to end users?
  - § Database researchers usually ignore this one!
- § Is uncertainty (in databases) only useful with large amounts of data?
  - § Can we scale?
    - § What is large scale?