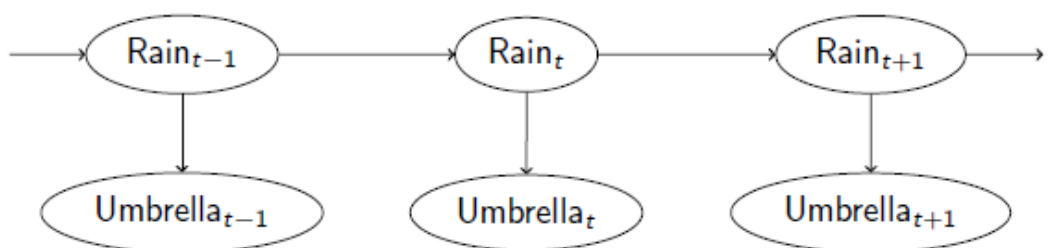


## Gedächtnisprotokoll Pattern Analysis

- Prüfer: Dr.-Ing. Christian Rieß
- Zweitprüfer: hat sich nicht vorgestellt
- Uhrzeit& Ort: kamen am Tag vorher per Mail
- Stimmung: entspannt
- Benotung: mit PR vergleichbar 😊
- Zeit: verging wie im Fluge

1. Übersicht: PA-Wolke malen (siehe 2017-SS\_PA\_Summary.pdf)
2. Random Forests:
  - a. Wie funktioniert unsupervised learning mittels Density Estimation Tree?
  - b. Wie schaut ein Regression tree aus?
3. Density Estimation:
  - a. Parzen-Window-Estimator: Formel für  $p(x)$  herleiten
  - b. Welche Alternative gibt es für die Density Estimation (s. Frage 2a)
4. Clustering:
  - a. Unterschiede zwischen k-means und mean-shift-algo (MSA) nennen, beide Algorithmen skizzieren
  - b. MSA:
    - i. Formel herleiten
    - ii. Welches Kernel verwendet man hier gerne? (Epanechnikov)
    - iii. Was passiert, wenn man nicht die 2-Norm, sondern eine andere Norm verwendet?
  - c. K-means:
    - i. Wie findet man ein gutes k? Mittels gap statistics (within-cluster-distance). Hier war es ihm besonders wichtig, dass die  $W(C)$ -Kurve streng monoton fallend ist, auch bei dem „little dent“. Da wäre es in der Vorlesung von Vorteil gewesen, die Tafelzeichnung korrekt abzumalen.
    - ii. Was passiert bei Outlayers (Punkte, die weit weg vom Rest vom Fest liegen)?
5. HMM:
  - a. HMM malen (Grafik aus KI2 geklaut)



- b. Beide Assumptions (Markov, Output) hinschreiben
  - c.  $p(x_1, \dots, x_N, y_1, \dots, y_N)$  vereinfachen: Indizes durch  $i \in [1; N]$  ersetzen, Bayes' Theorem anwenden
  - d. Welche Algorithmen gibt es? (nur die Namen haben gereicht)
6. Manifold Learning: dazu kam leider keine Frage, obwohl MDS und ISOMAP so schön zu erklären gewesen wären 😞