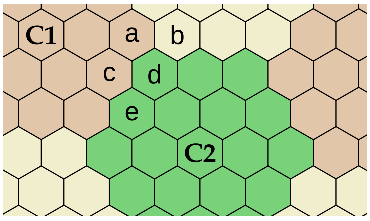
Composable Markov Building Blocks

Sander Evers everss@cs.utwente.nl
PhD student, Department of EEMCS, University of Twente



Some clusters (C1, C2) of granules (a, b, . . .)

In sensor data management, it is often useful to collect statistics about state changes of an observed phenomenon. The most basic case is to keep track of *how often certain transitions, e.g. from state* c *to* d*, have occurred*. In our research, we consider a discrete state space and timeline. For example, in the above figure, the state space corresponds to actual two-dimensional space and is divided into discrete *granules*. While an object moves through this space, its position (state) in terms of these granules is known at each point in time.

Counting the transitions between each pair of adjacent granules in a certain stretch of time (including self-transitions, points in time where the object stays put) provides exactly the information needed to construct a *first-order Markov model* of the object's behaviour. This model can be used to predict future behaviour or fill in missing readings.

However, in our research we assume that the state space is divided into clusters, and that it is impossible to track an object across a cluster boundary, due to technical or organizational issues. For example:

- Each cluster represents the area observed by one camera. By analyzing the difference between subsequent images, the camera can track moving objects, but cameras cannot send video to each other due to communication constraints. They can only keep count of transitions among the granules of their own area, and cluster exits/entrances at the border granules.
- Each cluster represents an area covered by one communication network. Due to different standards or privacy issues, people use a different ID when communicating with different networks.

Our goal is to deduce the missing statistics at the cluster boundaries by looking at the cluster exit and entrance counts of the different clusters. Thus, we compose a global Markov model.

As an illustration, consider the transition count $C(\mathsf{c}, \mathsf{d})$ between c and d. Cluster C1 can only count exit transitions $(\mathsf{c}, \mathsf{away}_1)$ and cluster C2 only entrance transitions $(\mathsf{away}_2, \mathsf{d})$. Nevertheless, from the arrangement of the granules it follows that

$$\begin{aligned} C(\mathsf{c}, \mathsf{away}_1) &= C(\mathsf{c}, \mathsf{d}) + C(\mathsf{c}, \mathsf{e}) \\ C(\mathsf{away}_2, \mathsf{d}) &= C(\mathsf{c}, \mathsf{d}) + C(\mathsf{a}, \mathsf{d}) + C(\mathsf{b}, \mathsf{d}) \end{aligned}$$

which, together with other transition counts, gives us enough information to deduce $C(\mathsf{c}, \mathsf{d})$. Previously[1], we have shown that under quite restrictive conditions on the inter-cluster transition graph, these deductions are exact. Currently, we present a method that broadens these conditions by approximating the answer. In the example with hexagonal clusters, the quality of this approximation is determined by the cluster radius. With a radius of 11, the approximation error is about 4%.

References

[1] Evers, S., Fokkinga, M.M., Apers, P.M.G.: Composable Markov Building Blocks. In Prade, H., Subrahmanian, V.S., eds.: Proceedings of the 1st International Conference on Scalable Uncertainty Management (SUM 2007), Washington DC, USA. Volume 4772 of LNCS., Berlin, Springer Verlag (October 2007) 131–142