

Summary of “Learning Systems of Concepts with an Infinite Relational Model”

Matt Dickenson mcd31

STA571/CS590.01

Due: 12 February, 2014

Kemp et al. (2006) introduce the Infinite Relational Model (IRM). The IRM is an unsupervised nonparametric Bayesian model that “discovers” which of a given set of entities belong to each of a (learned) number of subsets, and the relations between them. The prior on the number of groups is set according to the Chinese Restaurant Process (CRP), which helps the model to favor parsimony. From the input, which consists of types (e.g. people) and relations (e.g. “ i is friends with j ”), the model discovers a partition on the entities that sorts them into groups, each of which either mostly shares (1’s, in the binary case) or does not share (0’s in the binary case) the given relation. This discovered matrix, which is sorted by group, is assumed to have been generated from a partition z (on which the CRP prior is set) and a parameter matrix $\eta|\beta \sim \text{Beta}(\beta, \beta)$. The authors demonstrate the IRM on a variety of synthetic and real datasets, including animals, medical ontologies, aboriginal kinship relations, and international relations. Compared to an infinite mixture model (IMM), IRM has better clustering accuracy as measured by Rand indices.

From these examples, the IRM appears to do well at learning from binary relations between entities. The authors state that the IRM can be extended to discrete or continuous relations, but do not provide examples. One advantage of the IRM is that unlike some mixture models it does not require the number of clusters to be pre-specified, as discussed above. The IRM can also handle arbitrary collections of relations, making it more general than many models that are tailored for specific types of relations (e.g. a group-topic model that clusters people and words). A notable disadvantage of IRM is that it does not handle certain types of relational structures, such as hierarchies, as well as other social network models such as the stochastic block model. It also seems as if the IRM would handle asymmetric relations poorly. A final component missing from the paper is an indication of the computational complexity of IRM. The authors state that the model improves as more data is supplied, but do not address its capabilities for online learning. This model’s generality makes it attractive for a number of applications, but its accuracy may come at the expense of computational efficiency.