Math 269 - Combinatorics

Glenn Tesler
UCSD

Multi de Bruijn Sequences

Abstract:

We generalize the notion of a de Bruijn sequence to a “multi de Bruijn sequence”: a cyclic or linear sequence that contains every $k$-mer over an alphabet of size $q$ exactly $m$ times. For example, over the binary alphabet $\{0,1\}$, the cyclic sequence $(00010111)$ and the linear sequence $000101110$ each contain two instances of each 2-mer $00, 01, 10, 11$. We derive formulas for the number of such sequences. The formulas and derivation generalize classical de Bruijn sequences (the case $m = 1$). We also determine the number of multisets of aperiodic cyclic sequences containing every $k$-mer exactly $m$ times; for example, the pair of cyclic sequences $(00011)(011)$ contains two instances of each 2-mer listed above. This uses an extension of the Burrows-Wheeler Transform due to Mantaci et al., and generalizes a result by Higgins for the case $m = 1$.

Organizer: Jeff Remmel

Tuesday, April 11, 2017
4:00 PM
AP&M 7421