Generalization and Improvement to PPM's "Blending"

Suzanne Bunton

University of Washington

The best-performing method in the data compression literature for computing probability estimates of sequences on-line using a suffix-tree model is the *blending* technique used by PPM [CW84, Mof90]. Blending can be viewed as a bottom-up recursive procedure for computing a *mixture*, barring one missing term for each level of the recursion, where a mixture is basically a weighted average of several probability estimates. In [Bun97] we have shown by decomposition into an *inheritance weight* $in\{A, B, C, D\}$ and an *inheritance evaluation time*, M_h, that mixtures generalize the techniques used in DMC variants [CH87], as well as PPM variants, and thus these techniques, along with other variants of mixtures, are interchangeable.

Table 1 shows the relative effectiveness of most combinations of mixture weighting functions and inheritance evaluation times. Table 2 is a study on the value of using update exclusion, especially in models using state selection.

Table 1: How average compression performance on the Calgary Corpus as a whole is affected by varying mixture inheritance times and mixture weight functions, in models with and without (percolating) state selection.

Inherit Time	*A*9X*	*B*9X*	*C*9X*	*D*9X*	*A*9XS*	*B*9XS*	*C*9XS*	*D*9XS*
M_0	2.965	4.695	2.825	2.767	2.602	2.925	2.522	2.508
M_2	**2.631**	**2.505**	2.365	2.306	**2.386**	2.674	2.227	2.206
M_3	2.767	2.688	2.329	**2.281**	2.475	2.559	**2.197**	**2.191**
M_5	2.851	2.520	**2.300**	2.302	2.482	**2.419**	2.203	2.238

Table 2: The percent improvement of models using update exclusion over the same model variants without update exclusion.

Inherit Time	*A*9X*	*B*9X*	*C*9X*	*D*9X*	*A*9XS*	*B*9XS*	*C*9XS*	*D*9XS*
M_0	1.8	-6.5	1.1	2.9	2.0	0.2	6.9	2.1
M_2	5.5	3.2	2.2	5.9	5.7	3.7	3.2	5.4
M_3	5.9	-0.8	2.9	7.5	5.5	2.7	3.8	6.5
M_5	5.8	5.3	4.3	10.2	1.4	6.2	4.3	6.4

References

[Bun97] S. Bunton. A generalization and improvement to PPM's *blending*. EECS Technical Report 1.10, The University of Washington, January 1997.

[CH87] G. V. Cormack and R. N. S. Horspool. Data compression using dynamic Markov modelling. *The Computer Journal*, 30(6):541–550, 1987.

[CW84] J. G. Cleary and I. H. Witten. Data compression using adaptive coding and partial string matching. *IEEE Transactions on Communications*, 32(4):396–402, 1984.

[Mof90] A. Moffat. Implementing the PPM data compression scheme. *IEEE Transactions on Communications*, 38(11):1917–1921, 1990.

1068-0314/97$10.00 © 1997 IEEE