

Bibliography

- [1] ANSI/ISO C++ standard, November 1997. Electronic version of the standard available at <http://www.ncits.org>.
- [2] APPELT, D. E., AND MARTIN, D. Named entity extraction from speech: Approach and results using the TextPro system. In *Proceedings of the DARPA Broadcast News Workshop (HUB-4)* (February 1999).
- [3] BERGER, A., AND PRINTZ, H. A comparison of criteria for maximum entropy/minimum divergence feature selection. In *Proceedings of the Third Conference on Empirical Methods in Natural Language Processing* (June 1998), N. Ide and A. Boutilainen, Eds., The Association for Computational Linguistics, pp. 97–106.
- [4] BERGER, A. L., DELLA PIETRA, S. A., AND DELLA PIETRA, V. J. A maximum entropy approach to natural language processing. *Computational Linguistics* 22, 1 (1996), 39–71.
- [5] BIKEL, D. M., MILLER, S., SCHWARTZ, R., AND WEISCHEDEL, R. Nymble: a high-performance learning name-finder. In *Fifth Conference on Applied Natural Language Processing* (1997).
- [6] BLACK, W. J., RINALDI, F., AND MOWATT, D. Facile: Description of the NE system used for MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (1998).
- [7] BORTHWICK, A., FLORIAN, R., AND PAPINENI, K., 1997. MENE’s history, future, and feature base classes were influenced by the architecture of a maximum entropy language model jointly developed by these three authors at IBM Watson Labs, Yorktown Heights, New York, in the summer of 1997.
- [8] BORTHWICK, A., STERLING, J., AGICHTEIN, E., AND GRISHMAN, R. Exploiting diverse knowledge sources via maximum entropy in named entity recognition. In *Proceedings of the Sixth Workshop on Very Large Corpora* (August 1998).

- [9] BORTHWICK, A., STERLING, J., AGICHTEIN, E., AND GRISHMAN, R. NYU: Description of the MENE named entity system as used in MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (1998).
- [10] BROWN, P., DELLA PIETRA, S., DELLA PIETRA, V., MERCER, R., NADAS, A., AND ROUKOS, S. A maximum penalized entropy construction of conditional log-linear language and translation models using learned features and a generalized Csiszar algorithm. Unpublished IBM research report.
- [11] CHINCHOR, N. MUC-7 scoring methodology. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (April 1998).
- [12] CHINCHOR, N., BROWN, E., FERRO, L., AND ROBINSON, P. 1999 named entity recognition task definition, version 1.1 [draft], July 1999. Draft proposal for the 1999 HUB-4 named entity task.
- [13] CHINCHOR, N., BROWN, E., AND ROBINSON, P. Event99 named entity task definition, version 5.0 [draft], April 1999. Draft proposal for named entity task. This proposal was later substantially revised.
- [14] CHINCHOR, N., AND MARSH, E. MUC-7 named entity task definition. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (1998). Available at <http://www.muc.saic.com/>.
- [15] CORMEN, T. H., LEISERSON, C. E., AND RIVEST, R. L. *Introduction to Algorithms*. MIT Press, 1990, ch. 36.
- [16] COVER, T. M., AND THOMAS, J. A. *Elements of Information Theory*. John Wiley & Sons, Inc., New York, 1991, ch. 11.
- [17] CRYSTAL, M. R., AND KUBALA, F. Studies in data annotation effectiveness. In *Proceedings of the DARPA Broadcast News Workshop (HUB-4)* (February 1999), vol. 1.
- [18] DARROCH, J., AND RATCLIFF, D. Generalized iterative scaling for log-linear models. *The Annals of Mathematical Statistics* 43 (1972), 1470–1480.
- [19] DAVIS, E. Lecture for Introduction to Artificial Intelligence Class, New York University.
- [20] DELLA PIETRA, S., DELLA PIETRA, V., AND LAFFERTY, J. Inducing features of random fields. Tech. Rep. CMU-CS-95-144, Carnegie Mellon University, 1995.

- [21] ERIGUCHI, Y., AND KITANI, Y. NTT Data: Description of the Erie system used for MET-2 Japanese. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (1998).
- [22] FUKUMOTO, J., MASUMI, F., SHIMOHATA, M., AND SASAKI, M. Oki Electric Industry: Description of the Oki system as used for MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (April 1998). Available at <http://www.muc.saic.com/>.
- [23] FUKUMOTO, J., AND SHIMOHATA, M., 1999. Personal communication from J. Fukumoto and M. Shimohata.
- [24] GOOD, I. J. Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables. *Annals of Mathematical Statistics* 34 (1963), 911–934.
- [25] GRISHMAN, R. The NYU system for MUC-6 or where’s the syntax? In *Proceedings of the Sixth Message Understanding Conference* (November 1995), Morgan Kaufmann.
- [26] JAYNES, E. T. Information theory and statistical mechanics. *Physics Reviews* 106 (1957), 620–630.
- [27] JAYNES, E. T. Notes on present status and future prospects. In *Maximum Entropy and Bayesian Methods* (1990), W. T. Grandy and L. H. Schick, Eds., Kluwer, pp. 1–13.
- [28] JAYNES, E. T. Probability theory: The logic of science. Manuscript for book. Available at <http://bayes.wustl.edu/etj/prob.html>, July 1998.
- [29] KRUPKA, G. R., AND HAUSMAN, K. IsoQuest: Description of the NetOwl(tm) extractor system as used in MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (1998).
- [30] LIN, D. Using collocation statistics in information extraction. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (1998).
- [31] MAGERMAN, D. M. *Natural Language Parsing as Statistical Pattern Recognition*. PhD thesis, Stanford University, 1994.
- [32] MARSH, E., AND PERZANOWSKI, D. MUC-7 evaluation of I.E. technology: Overview of results. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (1998).
- [33] MASUI, F., 1998. Personal communication from Fumito Masui.

- [34] MATUMOTO, Y., KUROHASHI, S., YAMAJI, O., TAEKI, Y., AND NAGAO, M. Japanese morphological analyzing system: Juman. Kyoto University and Nara Institute of Science and Technology, 1997.
- [35] MIKHEEV, A., AND GROVER, C. LTG: Description of the NE recognition system used for MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (April 1998).
- [36] MILLER, D., SCHWARTZ, R., WEISCHEDEL, R., AND STONE, R. Named entity extraction from broadcast news. In *DARPA Broadcast News Workshop (HUB-4)* (February 1999).
- [37] MILLER, S., CRYSTAL, M., FOX, H., RAMSHAW, L., SCHWARTZ, R., STONE, R., WEISCHEDEL, R., AND THE ANNOTATION GROUP. Algorithms that learn to extract information—BBN: Description of the SIFT system as used for MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)* (April 1998).
- [38] PRINTZ, H. Fast computation of maximum entropy/minimum divergence model feature gain. In *Proceedings of the Fifth International Conference on Spoken Language Processing* (November 1998).
- [39] QUINLAN, R. J. *C4.5: Program for Machine Learning*. Morgan Kaufman Publishers, 1993.
- [40] RATNAPARKHI, A. A maximum entropy model for part-of-speech tagging. In *Conference on Empirical Methods in Natural Language Processing* (May 1996), University of Pennsylvania, pp. 133–142.
- [41] REYNAR, J. C., AND RATNAPARKHI, A. A maximum entropy approach to identifying sentence boundaries. In *Fifth Conference on Applied Natural Language Processing* (April 1997), pp. 16–19.
- [42] RISTAD, E. S. Maximum entropy modeling for discrete domains. Tutorial given at the M3D Workshop, University of Illinois at Chicago.
- [43] RISTAD, E. S. Maximum entropy modeling toolkit, release 1.6 beta, February 1998. Includes documentation which has an overview of MaxEnt modeling.
- [44] ROSENFELD, R. *Adaptive Statistical Language Modeling: A Maximum Entropy Approach*. PhD thesis, Carnegie Mellon University, 1994. CMU Technical Report CMU-CS-94-138.
- [45] SEKINE, S., 1998. Personal communication from Satoshi Sekine.