MEXT Grant-in-Aid for Scientific Research on Priority Areas
"Statistical Mechanical Approach to Probabilistic Information Processing"
Workshop on Mathematics of Statistical Inference (December 2004, Tohoku University, Sendai, Japan)

Preface

Kazuyuki Tanaka [1]

Graduate School of Information Sciences, Tohoku University,

Aramaki-aza-aoba 09, Aoba-ku, Sendai 980-8579, Japan

Yoshiyuki Kabashima

Department of Computational Intelligence and Systems Science,

Tokyo Institute of Technology, Yokohama 226-8502, Japan

Toshiyuki Tanaka

Department of Electronics and Information Engineering,

Tokyo Metropolitan University, Tokyo 192-0397, Japan

This workshop, held at the Aoba Memorial Hall of Tohoku University, Sendai, Japan, from 6 to 8 December, 2004, is organized to survey recent advances in theoretical studies of algorithms for statistical inference, with emphasis on the belief propagation and related algorithms, and to discuss future directions of research. We believe that belief propagation and related algorithms will play a significant role in establishing a new computational paradigm in the twenty-first century, and hope this workshop to help form a basis for such a development.

What makes studies of the belief propagation so important, we think, is twofold: One is lack of thorough theory, and another is its interdisciplinarity. The belief propagation was proposed by Pearl [1] in mid 1980s, as an efficient algorithm of statistical inference, which aims at implementing probability-based artificial intelligence capable of handling uncertainties. The belief propagation is proved to give exact results in the special cases where the graphical representation of the underlying probability model has no loops. However, what is interesting and important is beyond that. Although the belief propagation is not guaranteed to give exact results for general graphical models with loops, it works reasonably well in various cases, and furthermore, in some specific cases it turns out to work excellently. The most ground-breaking example in this respect appeared in the field of coding theory. It was the turbo codes, proposed by Berrou and Glavieux [2], the decoding algorithm of which was later shown by McEliece, MacKay, and Cheng [3] as nothing but the belief propagation applied to probability models with loops. The turbo decoding has shown surprisingly high performance despite that it lacks theoretical justification of giving good decoding results, which ignited intense research activities in the field of coding theory. Soon after that came the low-density parity-check (LDPC) codes, which have been rediscovered by MacKay

[1]E-mail: kazu@statp.is.tohoku.ac.jp

and Neal [4] after thirty years since the original proposal by Gallager [5], which also exhibited high performance. Interestingly enough, the sum-product algorithm for decoding of the LDPC codes again turns out to be the belief propagation. It has now been widely accepted in the coding theory community that iterative decoding algorithms based on the belief propagation is the key to the high performance of these codes.

The belief propagation also has a link with statistical mechanics. Kabashima and Saad [6], in their study of constructing a decoding algorithm of one type of the LDPC codes based on statistical mechanical framework, were the first to mention the link, by showing that the Bethe approximation, which is one of the advanced mean-field methods in statistical mechanics [7], yields the belief propagation when it is applied to a system describing the decoding problem. Inspired by their work, Yedidia, Freeman, and Weiss [8] showed that the link is indeed general, that is, they established correspondence between the belief propagation and the Bethe approximation in general settings. Their work has indeed opened a new stage of research. In fact, mean-field methods have a long history in statistical mechanics, and several notions and tools have been developed to better understand properties of a system with many interacting elements. It has become possible to try using these notions and tools in order to obtain better understanding of the belief propagation and related algorithms. For example, the Bethe free energy, whose stationary points give self-consistent conditions of the Bethe approximation, has been introduced [8] and then extensively studied to characterize stationary states of the belief propagation. Based on the introduction of the Bethe free energy, Yuille [9] proposed the concave-convex procedure (CCCP) as an alternative algorithm to minimizing the Bethe free energy. Another important example would be the cluster variation method [10, 11, 12, 13], one of yet advanced mean-field methods. Yedidia et al. [8] proposed the generalized belief propagation on the basis of the cluster variation method, while Tanaka and Morita, independent of Yedidia et al.'s proposal and long before that, applied the framework of the cluster variation method to problems of statistical image restoration [14]. In recent years, applications of these advanced mean-field approaches to various information processing problems are indeed very active research fields, including decoding of error control codes [15], user detection in code-division multiple-access communications [16, 17], image processing [14, 18, 19, 20], solving optimization problems [21, 22], inference in Bayesian networks [23, 24], machine learning [25, 26, 27, 28], to mention a few. These examples suggest that statistical mechanics, in particular the advanced mean-field framework, is a vein of various interesting and inspiring ideas for studying the belief propagation and related algorithms, as well as their applications [7, 29].

Despite these achievements, we feel that we are still far from complete understanding of the belief propagation as an algorithm of statistical inference. Important problems are the accuracy of the algorithm's result and its dynamical properties, when the algorithm is applied to problems with loops. Several ideas have been proposed to study these problems. They include: the density evolution framework [30] to describe macroscopic dynamics of the belief propagation applied to a system with uniform randomness, Weiss's idea of loop unwrapping [31, 32] and the notion of the limiting Gibbs measures [33], the information-geometrical interpretation of the belief propagation and related algorithms [34, 35], and so on. The problems and approaches to them span a wide variety of interdisciplinary fields across mathematics, physics, and engineering, including mathematical statistics, differential geometry, spin-glass theory, information theory, coding theory, computer science, artificial intelligence, pattern recognition, neural networks, etc. etc. This view has motivated us to organize this workshop, expecting that it will further facilitate interdisciplinary collaboration, which in turn will put forward the research uncovering the secrets of the belief propagation and related algorithms.

The workshop is entitled "Mathematics of Statistical Inference," and is organized as a workshop of the MEXT Grant-in-Aid for Scientific Research on Priority Areas: Statistical-Mechanical Approach to Probabilistic Information Processing (SMAPIP). This workshop covers the following topics:

1. Belief propagation and advanced mean-field methods
2. Approximate statistical inference
3. Spin glass theory and statistical inference

The organizing committee of this workshop consists of:

Kazuyuki Tanaka, SMAPIP Head Investigator (Tohoku University)

Yoshiyuki Kabashima (Tokyo Institute of Technology)

Toshiyuki Tanaka (Tokyo Metropolitan University)

We would like to thank all participants and hope that this workshop will be very successful and will contribute to the developments of this strongly interdisciplinary field.

References

[1] J. Pearl: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988.

[2] C. Berrou and A. Glavieux: Near optimum error correcting coding and decoding: Turbo-codes, IEEE Trans. Comm., vol. 44, pp. 1261–1271, 1996.

[3] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng: Turbo decoding as an instance of Pearl's 'belief propagation' algorithm, IEEE J. Sel. Areas Comm., vol. 16, no. 2, pp. 140–152, 1998.

[4] D. J. C. MacKay and R. M. Neal: Near Shannon limit performance of low density parity check codes, Electron. Lett., vol. 32, no. 18, pp. 1645–1646, 1996; Reprinted *ibid.*, vol. 33, no. 6, pp. 457–458, 1997.

[5] R. G. Gallager: Low density parity check codes, IRE Trans. Info. Theory, vol. IT-8, pp. 21–28, 1962.

[6] Y. Kabashima and D. Saad: Belief propagation *vs.* TAP for decoding corrupted messages, Europhysics Letters, vol. 44, no. 5, pp. 668–674, Dec. 1998.

[7] M. Opper and D. Saad (eds): *Advanced Mean Field Methods — Theory and Practice —*, MIT Press, 2001.

[8] J. S. Yedidia, W. T. Freeman and Y. Weiss: Generalized belief propagation, Advances in Neural Information Processing Systems, vol. 13, pp. 689–695, MIT Press, 2001.

[9] A. L. Yuille: CCCP algorithms to minimize the Bethe and Kikuchi free energies: convergent alternatives to belief propagation, Neural Computation, vol. 14, no. 7, pp. 1691–1722, 2002.

[10] R. Kikuchi: A theory of cooperative phenomena, Phys. Rev., vol. 81, no. 6, pp. 988–1003, March 1951.

[11] T. Morita: General structure of the distribution functions for the Heisenberg model and the Ising model, J. Math. Phys., vol. 13, no. 1, pp. 115–123, 1972.

[12] T. Horiguchi: On the Bethe approximation for the random bond Ising model, Physica A, vol. 107, no. 2, pp. 360–370, 1981.

[13] T. Morita, M. Suzuki, K. Wada and M. Kaburagi (eds): *Foundations and Applications of Cluster Variation Method and Path Probability Method*, Prog. Theor. Phys., Supplement, no. 115, 1994.

[14] K. Tanaka and T. Morita: Cluster variation method and image restoration problem. Phys. Lett. A, vol. 203, nos. 2–3, pp. 122–128, 1995.

[15] Y. Kabashima and D. Saad: Statistial mechanics of low-density parity-check codes (Topical Review), J. Phys. A: Math. & Gen., vol. 37, no. 6, pp. R1–R43, 2004.

[16] Y. Kabashima: A CDMA multiuser detection algorithm on the basis of belief propagation, J. Phys. A: Math. & Gen., vol. 36, no. 43, pp. 11111–11121, Oct. 2003.

[17] T. Tanaka and M. Okada: Approximate belief propagation, density evolution, and statistical neurodynamics for CDMA multiuser detection, to appear in IEEE Trans. Inform. Theory, 2004.

[18] K. Tanaka: Statistical-mechanical approach to image processing (Topical Review), J. Phys. A: Math. & Gen., vol. 35, no. 37, pp. R81–R150, 2002.

[19] W. T. Freeman, T. R. Jones and E. C. Pasztor, Example-based super-resolution, IEEE Computer Graphics and Applications, vol. 22, no. 2, pp. 56–65, 2002.

[20] K. Tanaka, H. Shouno, M. Okada and D. M. Titterington: Accuracy of the Bethe Approximation for Hyperparameter Estimation in Probabilistic Image Processing, J. Phys. A: Math. & Gen., vol. 37, no. 36, pp. 8675–8696, 2004.

[21] O. C. Martin, R. Monasson, R. Zecchina: Statistical mechanics methods and phase transitions in optimization problems, Theoretical Computer Science, vol. 265, pp. 3–67, 2001.

[22] M. Mezard, G. Parisi, R. Zecchina: Analytic and algorithmic solution of random satisfability problems, Science, vol. 297, no. 1/2, pp. 812–815, 2002.

[23] H. J. Kappen and W. Wiegerinck: Novel iteration schemes for the cluster variation method, Advances in Neural Information Processing System, vol. 14, pp. 415–422, MIT Press, 2002.

[24] K. Tanaka: Probabilistic inference by means of cluster variation method and linear response theory, IEICE Transactions on Information and Systems, vol. E86-D, no. 7, pp. 1228–1242, 2003.

[25] H. J. Kappen and F. B. Rodríguez: Efficient learning in Boltzmann machines using linear response theory, Neural Computation, vol. 10, no. 5, pp. 1137–1156, 1998.

[26] T. Tanaka: Mean field theory of Boltzmann machine learning, Phys. Rev. E, vol. 58, no. 2, pp. 2302-2310, 1998.

[27] M. Welling and Y. W. Teh: Approximate inference in Boltzmann machines, Artificial Intelligence, vol. 143, pp. 19–50, 2003.

[28] M. Welling and Y. W. Teh: Linear response algorithms for approximate inference in graphical models, Neural Computation, vol. 16, no. 1, pp. 197–221, Jan. 2004.

[29] H. Nishimori: *Statistical Physics of Spin Glasses and Information Processing: An Introduction*, Oxford University Press, Oxford, 2001.

[30] T. J. Richardson and R. L. Urbanke: The capacity of low-density parity-check codes under message-passing decoding, IEEE Trans. Inform. Theory, vol. 47, no. 2, pp. 599–618, 2001.

[31] Y. Weiss: Correctness of local probability propagation in graphical models with loops, Neural Computation, vol. 12, no. 1, pp. 1–41, 2000.

[32] Y. Weiss and W. T. Freeman: Correctness of belief propagation in Gaussian graphical models of arbitrary topology, Neural Computation, vol. 13, no. 10, pp. 2173–2200, 2001.

[33] S. Tatikonda and M. I. Jordan: Loopy belief propagation and Gibbs measures, D. Koller and A. Darwiche (eds.), Uncertainty in Artificial Intelligence, Proc. 18th Conf., pp. 493–500, 2002.

[34] S. Ikeda, T. Tanaka and S. Amari: Information geometry of turbo and low-density parity-check codes, IEEE Transactions on Information Theory, vol. 50, no. 6, pp. 1097–1114, 2004.

[35] S. Ikeda, T. Tanaka and S. Amari: Stochastic reasoning, free energy, and information geometry, Neural Computation, vol. 16, no. 9, pp. 1779–1810, 2004.