



# HOKKAIDO UNIVERSITY

Title	Asymptotics of Discrete MDL for Online Prediction
Author(s)	Poland, Jan; Hutter, Marcus
Description	Erratum published: IEEE Transactions on Information Theory. Volume 52, Issue 3, March 2006, p.1279 <a href="https://doi.org/10.1109/TIT.2006.869753">https://doi.org/10.1109/TIT.2006.869753</a>
Citation	IEEE Transactions on Information Theory, 51(11), 3780-3795 <a href="https://doi.org/10.1109/TIT.2005.856956">https://doi.org/10.1109/TIT.2005.856956</a>
Issue Date	2005-11
Doc URL	<a href="https://hdl.handle.net/2115/8468">https://hdl.handle.net/2115/8468</a>
Rights	(c) 2005-2006 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
Type	journal article
File Information	01603798.pdf, erratum



# Erratum

---

## Erratum to “Asymptotics of Discrete MDL for Online Prediction”

Jan Poland

The previously published abstract [1], erroneously stated that the work is about learning i.i.d. processes. But in fact, the main contribution are methods and proof techniques which work with arbitrary processes on a finite observation space. We regret any misunderstanding this might have caused.

The corrected version of the Abstract should read as follows:

**Abstract**—Minimum description length (MDL) is an important principle for induction and prediction, with strong relations to optimal Bayesian learning. This paper deals with learning processes which are not necessarily independent and identically distributed, by means of two-part MDL, where the underlying model class is countable. We consider the online learning framework, i.e., observations come in one by one, and the predictor is allowed to update its state of mind after each time step. We identify two ways of predicting by MDL for this setup, namely, a static and a dynamic one. (A third variant, hybrid MDL, will turn out inferior.) We will prove that under the only assumption that the data is generated by a distribution contained in the model class, the MDL predictions converge to the true values almost surely. This is accomplished by proving finite bounds on the quadratic, the Hellinger, and the KullbackLeibler loss of the MDL learner, which are, however, exponentially worse than for Bayesian prediction. We demonstrate that these bounds are sharp, even for model classes containing only Bernoulli distributions. We show how these bounds imply regret bounds for arbitrary loss functions. Our results apply to a wide range of setups, namely, sequence prediction, pattern classification, regression, and universal induction in the sense of algorithmic information theory among others.

### REFERENCES

- [1] J. Poland and M. Hutter, “Asymptotics of Discrete MDL for Online Prediction,” *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3780–3795, Nov. 2005.

Manuscript received December 26, 2005.

J. Poland is with the Graduate School of Information Science and Technology, Hokkaido University, N-14, W-9 Sapporo 060-0814, Japan (e-mail: jan@ist.hokudai.ac.jp).

Digital Object Identifier 10.1109/TIT.2006.869753