

This bibliography is by no means complete or exhaustive about research works linked to Training Issues in Incremental Learning. It is only intended as a starting point for new researchers in the field, and, for others, as hints for references. It reflects of course the current state of the art, and it is expected _ and hoped _, that it will shortly be superseded.

- John ANDERSON & Michael MATESSA (1992) : **"Explorations of an Incremental, Bayesian Algorithm for Categorization"**. *Machine Learning*, 9, pp.275-308, (1992).
- Dana ANGLUIN & Carl SMITH (1983) : **"Inductive Inference: Theory and Methods"**. *Computing Surveys*, Vol.15, No.3, 237-269, September 1983.
A survey paper providing providing a broad outline of definitions, key notions and known results in 1983. Discuss several criteria for comparing inference methods including the number of distinct hypotheses and mind changes produced by an inference system given a sequence of examples. States (p.254) that "In general, identification by enumeration produces the 'simplest' compatible hypothesis if the enumeration order is consistent with the intended simplicity ordering of hypotheses".
- Gyora BENEDEK & Alon ITAI (1988) : **"Learnability by fixed distributions"**. In *Proc. of COLT'88* (Workshop on Computational Learning Theory), Cambridge 1988, pp.80-90.
Proposes a theory of PAC-learning for particular distributions known in advance by the learner. Interesting if one sees supervised incremental learning as a modification of the prior distribution of the examples.
- BLUM L. & BLUM M. (1975) : **"Toward a mathematical theory of inductive inference"**. *Inf. Control* 28, 125-155.
Shows that, in the case of partial recursive functions, identification methods may be made order independent without loss of generality.
- Antoine CORNUEJOLS (1989a) : **"De l'Apprentissage Incrémental par Adaptation Dynamique: le système INFLUENCE"** (Incremental Learning through Dynamical Adaptation: the INFLUENCE system). Doctoral Dissertation. University of Paris-Sud, Orsay, France, January 6th 1989.
- Antoine CORNUEJOLS (1989b) : **"An Exploration into Incremental Learning : the INFLUENCE System"**, in *Proc. of the 6th Intl. Conf. on Machine Learning*, Ithaca, June 29- July 1, 1989, pp.383-386.
- Antoine CORNUEJOLS (1993) : **"Getting Order Independence in Incremental Learning"**, in *Proc. of the European Conf. on Machine Learning*, (ECML-93), Vienna, April 5-8, 1993.
Shows that it is not possible in general for an incremental learning system to have a strategy of forgetting hypotheses during learning and stay order independent. Provides an interesting means of translating conditions on order dependent training sequences into conditions on an order independent training set.

- Robert DALEY & Carl SMITH (1986) : **"On the Complexity of Inductive Inference"**. *Information and Control*, 69, pp.12-40, 1986.
Introduces an axiomatization of the concept of the complexity of inductive inference based on the total resources used by an inference strategy in the process of converging to a correct inference. (distinguihes it from other measures such as the number of distinct hypotheses made by the inference system or the number of mind change in the course of inferencing). Gives some examples of the effect of changing the order of presentation on the complexity of inference.
- Werner EMDE (1987) : **"Non-cumulative learning in METAXA.3"**. *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI-87)*, Milano, 1987, pp.208-210.
- Douglas FISHER (1987) : **"Knowledge Acquisition Via Incremental Conceptual Clustering"**, *Machine Learning*, 2, pp.139-172, 1987.
- FISHER, XU & ZARD (1992) : **"Ordering Effects in COBWEB and an Order-Independent Method"**. To appear in *Proc. of the 9th Int. Conf. on Machine Learning*, Aberdeen, june 29-july 1st, 1992.
- Sally FLOYD (1989) : **"Space-Bounded Learning and the Vapnik-Chervonenkis Dimension"**. In *Proc. of COLT-89*, 349-364.
- GENNARI, LANGLEY & FISHER (1989) : **"Models of Incremental Concept Formation"**. *Artificial Inteligence*, 40, 1989, pp.11-61.
- Sally GOLDMAN & Michael KEARNS (1991) : **"On the Complexity of Teaching"**. *Proc. of COLT'91*, Santa Cruz, Aug. 5-7 1991, pp.303-314.
Defines a Teaching Dimension (the minimum number of examples a techer must reveal to uniquely identify any concept in the class of concepts). Prove that finding an optimal teaching sequence is NP-hard. Give optimal teaching sequences for some classes of concepts.
- Mirsad HADZIKADIC & David YUN (1989) : **"Concept Formation by Incremental Conceptual Clustering"**, *Proc. of the IJCAI-89*, pp.831-836.
- HAUSSLER, KEARNS & SCHAPIRE (1991) : **"Bounds on the Sample Complexity of Bayesian Learning Using Information Theory and the VC Dimension"**. *Proc. of COLT'91*, Santa Cruz, Aug. 5-7 1991, pp.61-74.
- HELMBOLD David & LONG Philip (1991) : **"Tracking Drifting Concepts Using Random Examples"**. *Proc. of COLT'91*, Santa Cruz, Aug. 5-7 1991, pp.13-23.
Related to rate effects. What is the optimal rate with which to follow a drifting concept?
- HELMBOLD, SLOAN & WARMUTH (1989) : **"Learning Nested Differences of Intersection-Closed Concept Classes"**. In *Proc. of COLT-89*, 41-56.
- Haym HIRSH (1990) : **"Incremental Version-Space Merging"**. *Proc. of the 7th Int. Conf. on Machine Learning*. Univ. of Austin, Texas, June 21-23, 1990, pp.330-338.

- Haim HIRSH (1989) : **Incremental version-space merging : a general framework for concept learning**. Doctoral dissertation, Stanford, june 1989.
Says on p. 65 that "the general problem of determining a subset of training data that yields optimal learning is NP-hard" (!?). Provides an heuristic for an optimal ordering of the training examples.
- James MacGREGOR (1988) : **"The Effects of Order on Learning Classifications by Example: Heuristics for Finding the Optimal Order"**, *Artificial Intelligence*, vol.34, pp.361-370, 1988.
- IBA, WOGULIS & LANGLEY (1988) : **"Trading off simplicity and coverage in incremental concept learning"**. *Proc. Of the 5th Intl. Conf. on Machine Learning*, pp.73-79, Ann Arbor, Michigan, 1988. Morgan Kaufmann.
- Klaus JANTKE (1990) : **"Monotonic and non-monotonic Inductive Inference"**. *Proc. of the first Int. workshop on Algorithmic Learning Theory (ALT'90)*, Tokyo, October 8-10,1990, pp.269-281.
- LANGLEY, GENNARI & IBA (1987) : **"Hill-Climbing Theories of Learning"**. In *Proc. of the 4th Int. Workshop on Machine Learning*, Univ. of Calif., Irvine, June 22-25, 1987, pp.312-323.
Describes the CLASSIT and MAGGIE systems, both incremental and maintaining only one hypothesis at a time without memory of previous states.
- Xiaofeng (Charles) LING (1991) : **"Inductive Learning from Good Examples"**. In *Proc. of the 12th Int. Joint Conf. on Artif. Intel. (IJCAI-91)*, pp.751-756.
Studies what kind of data may ease the computational complexity of learning of Horn clause theories and boolean functions.
- Wolfgang MAASS (1991) : **"On-line Learning with an Oblivious Environment and the Power of Randomization"**. *Proc. of COLT'91*, Santa Cruz, Aug. 5-7 1991, pp.167-175.
- Wolfgang MAASS & György TURAN (1992) : **"Lower Bound Methods and Separation Results for On-Line Learning Models"**. *Machine Learning*, 9, pp.107-145, 1992.
- Tom MITCHELL (1978) : **Version Spaces : An Approach to Concept Learning**. Ph.D. Dissertation, Stanford University, december 1978, *STAN-CS-78-711,HPP-79-2*.
- Tom MITCHELL (1982) : **"Generalization as Search"**, *Artificial Intelligence*, vol.18, pp.203-226, 1982.
- Jordan B. POLLACK (1991) : **"The Induction of Dynamical Recognizers"**. *Machine Learning*, 7, pp.227-252, 1991.
Studies learning networks from the perspective of dynamical systems.
- Sara PORAT & Jerome A. FELDMAN (1991) : **"Learning Automata from Ordered Examples"**. *Machine Learning*,7, pp.109-138, 1991.
This paper presents a provably correct algorithm for inferring any minimum-state deterministic finite-state automata (FSA) from a complete ordered sample using limited total storage and

without storing examples strings. It also shows that no machine with finite working storage can iteratively identify the finite-state language (FSL) from arbitrary presentations.

- Anand RAO & Norman FOO (1989) : **"Minimal Change and Maximal Coherence: A Basis for Belief Revision and Reasoning about Actions"**. *Proc. of IJCAI-89*, 966-971.
- REINKE & MICHALSKI (1986) : **"Incremental Learning of Concept Description : A Method and Experimental Results"**. In J.E. Hayes (Ed.), *Machine Learning*, 11, Oxford University Press, 1986.
- Ronald RIVEST & Robert SLOAN (1988) : **"Learning Complicated Concepts Reliably and Usefully"**. In *Proc. of COLT'88* (Workshop on Computational Learning Theory), Cambridge 1988, pp.69-79.
Proposes to help the learner by having learn first sub-concepts of the concept to be learned. And defines a variation of the Valiant model by introducing the notions of reliable and useful learning.
- J. SCHLIMMER & D. FISHER (1986) : **"A Case Study of Incremental Concept Induction"**. *AAAI-86*, 1986, pp.496-501.
- J. SCHLIMMER & R. GRANGER (1986) : **"Incremental Learning from Noisy Data"**. *Machine Learning*, pp.317-354, 1986.
- J. SCHLIMMER & R. GRANGER (1986) : **"Beyond incremental processing: Tracking concept drift"**. *Proc. of the AAAI-86*, pp.502-507, 1986.
- SEUNG & SOMPOLINSKY & TISHBY (1991) : **"Learning Curves in Large Neural Networks"**. *Proc. of COLT'91*, Santa Cruz, Aug. 5-7 1991, pp.113-127.
- Benjamin SMITH & Paul ROSENBLOOM (1990) : **"Incremental Non-Backtracking Focusing: A polynomially Bounded Generalization Algorithm for Version Spaces"**. In *Proc. of the AAAI-90*, pp.848-853.
- Ayumi SHINOHARA & Satoru MIYANO (1990) : **"Teachability in Computational Learning"**, in *Proc. of the Workshop on Algorithmic Learning Theory*, 1990, pp.247-255.
Establishes a relationship between the notions of teachability and learnability.
- Devika SUBRAMANIAN & Joan FEIGENBAUM (1986) : **"Factorization in Experiment Generation"**. In *Proc. of the AAAI-86*, pp.518-522.
Proposes to speed learning in the Version Space scheme by exploiting independence properties and factoring the Version Space.
- Paul UTGOFF (1988) : **"ID5 : An Incremental ID3"**. In *Proc. of the 5th Intl. Conf. on Machine Learning*, Ann Arbor, June 12-14, 1988, pp.107-120.
- Paul UTGOFF (1989) : **"Incremental Induction of Decision Trees"**. *Machine Learning*, 4, pp.161-186, 1989.

- Van de Velde W. (1989) : **"IDL, or Taming the Multiplexer"**. *Proc. of EWSL-89*, Montpellier, France, 4-6 december 1989, pp.211-226.
- C. Vrain & C. Lu (1988) : **"An Analogical method to do incremental learning of concepts"**. *Proc. of the EWSL-88*, Glasgow, 1988.
- Kurt Van LEHN (1987) : **"Learning one subprocedure per lesson"**. *Artificial Intelligence*, 31 (1), pp.1-40, january 1987.
- Widmer G. (1989) : **"An Incremental Version of Bergadano & Giordana's Integrated Learning Strategy"**. *Proc. of EWSL-89*, Montpellier, France, 4-6 december 1989, pp.227-236.
- WINSTON P.H. (1970) : **"Learning structural descriptions from examples"**, AI-TR-231, MIT, Artificial Intelligence Laboratory, Cambridge, MA, 1970.