

## ANDREA PASSERINI

### Office Address

Dipartimento di Ingegneria e Scienza dell'Informazione  
Università degli Studi di Trento  
D10, Via di Sommarive 14  
POVO 38100 TN, Italy  
Tel: +39 0461 88 5224  
e-mail: [passerini@disi.unitn.it](mailto:passerini@disi.unitn.it)  
web page: <http://www.disi.unitn.it/~passerini>

### Personal information

**Andrea Passerini**

**Date of birth:** 27 March 1975

**Place of birth:** Firenze

**Citizenship:** Italian

### Education

DIPARTIMENTO DI SISTEMI E INFORMATICA  
UNIVERSITA' DEGLI STUDI DI FIRENZE April 2001–March 2004  
PhD in Computer Engineering.  
Advisors: Prof. Paolo Frasconi, Prof. Giovanni Soda.

- PhD Thesis  
*“Kernel Methods, Multiclass Classification and Applications to Computational Molecular Biology”*

PhD Title in Computer Engineering obtained in date 21.05.2004.

UNIVERSITA' DEGLI STUDI DI FIRENZE Master of Science in Computer Engineering  
in date 06.11.2000 with the voting of 110/110 Cum Laude.

- Master Thesis  
*“Tecniche di apprendimento automatico applicate al recupero di informazione da Internet”*  
Reporters: Prof. Paolo Frasconi, Prof. Giovanni Soda.

### Academic Positions

DIPARTIMENTO DI INGEGNERIA E SCIENZA DELL'INFORMAZIONE  
UNIVERSITA' DEGLI STUDI DI TRENTO  
September 2009–currently  
Assistant Professor.

DIPARTIMENTO DI INGEGNERIA E SCIENZA DELL'INFORMAZIONE  
UNIVERSITA' DEGLI STUDI DI TRENTO  
July 2008–August 2009  
Visiting Professor.

DIPARTIMENTO DI SISTEMI E INFORMATICA  
UNIVERSITA' DEGLI STUDI DI FIRENZE  
July 2007–June 2008

Assegno di ricerca: *Rappresentazioni logiche probabilistiche e metodi di inferenza per l'apprendimento in domini strutturati.*

DIPARTIMENTO DI SISTEMI E INFORMATICA  
UNIVERSITA' DEGLI STUDI DI FIRENZE  
July 2004–June 2007

Assegno di ricerca: *Rappresentazioni logiche probabilistiche e metodi di inferenza per l'apprendimento in domini strutturati*, within the FET APrIL II project (*Applications of Probabilistic Inductive Logic Programming*).

## Research projects

LEARNING TECHNIQUES IN RELATIONAL DOMAINS AND THEIR APPLICATIONS  
2011-2013  
PRIN. Role: Researcher.

A-CUBE: AMBIENT AWARE ASSISTANCE  
2008-2011  
“Grand Projects” funding from the Autonomous Province of Trento.  
Role: Researcher.

APRIL II: APPLICATIONS OF PROBABILISTIC INDUCTIVE LOGIC PROGRAMMING  
2004-2006  
STREP funded by the European Commission.  
Role: Researcher.

STRUMENTI BASATI SU APPRENDIMENTO AUTOMATICO PER LA GENOMICA STRUTTURALE E FUNZIONALE  
2002-2003  
PRIN.  
Role: Researcher.

## Main research activities

My research activity is primarily in the areas of machine learning, especially kernel machines, statistical logical learning, learning for optimization, and bioinformatics, dealing with protein function and structure prediction. I have published over 50 refereed papers, including 19 journal publications and three invited book chapters. My h-index is 12 (Google Scholar, Jan 2012). Below is a list of my main research achievements:

- *Bridging the gap between statistical and symbolic learning*: kernels on program traces [35][53][16]; kernels on logical representations [23][32][39, 10][48]; type extension trees for modelling combinatorial features [30][49]; relational information gain [7][28].
- *Machine learning algorithms for protein structure and function prediction*: prediction of disulphide bonding state and bridges [17][13][21][38]; prediction of metal binding sites [27][29][4, 5, 6, 12][14][33][18]; prediction of secondary [40] and tertiary [26][31] structure; prediction of catalytic residues in enzymes [9]; predicting mutations in viruses [44] [48].
- *Multiclass classification*: error-correcting output codes [37][19].
- *Learning complex data structures*: structured-output learning with probabilistic [11][51] and search-based [27][29] approaches.
- *Learning to optimize*: machine learning approaches for interactive decision making [8][25][45, 47].

## Research collaborations

GIANLUCA POLLASTRI  
 AMMBIO LAB  
 SCHOOL OF COMPUTER SCIENCE AND INFORMATICS  
 UNIVERSITY COLLEGE DUBLIN  
 DUBLIN, IRELAND  
 2007–currently

Ranking candidate protein three-dimensional structures using information from the pseudo-folding history of the simulation algorithm.  
 [31]

MANFRED JAEGER  
 INSTITUT FOR DATALOGI  
 AALBORG UNIVERSITET  
 AALBORG, DENMARK  
 2006–currently

Learning in relational domains using complex combinatorial features, relational information gain.  
 [30][49][7][28]

IVANO BERTINI  
 MAGNETIC RESONANCE CENTER (CERM)  
 DIPARTIMENTO DI CHIMICA  
 UNIVERSITA' DEGLI STUDI DI FIRENZE  
 SESTO FIORENTINO (FI), ITALIA  
 2005–2007

Learning algorithms and architectures for zinc binding sites prediction and mining at the proteome level.  
 [14][33]

STEPHEN MUGGLETON  
COMPUTATIONAL BIOINFORMATICS LABORATORY  
DEPARTMENT OF COMPUTING  
IMPERIAL COLLEGE LONDON  
LONDON, UK  
2005–2006

Kernels over structured objects using mereotopological relations.  
[54]

LUC DE RAEDT  
DECLARATIVE LANGUAGES AND ARTIFICIAL INTELLIGENCE  
DEPARTEMENT COMPUTERWETENSCHAPPEN  
KATHOLIEKE UNIVERSITEIT LEUVEN  
LEUVEN, BELGIUM  
2004–currently

Hybrid approaches combining statistical and relational learning  
[10][16][32][53]

BURKHARD ROST  
BIOINFORMATIC AND COMPUTATIONAL BIOLOGY DEPARTEMENT  
TECHNISCHE UNIVERSITAET MUENCHEN  
MUENCHEN, GERMANY  
2004–currently

Learning algorithms and architectures for metal binding sites prediction, and applications to target selection and function annotation within structural genomics projects.  
[6, 12, 18]

## Teaching experiences

UNIVERSITA' DEGLI STUDI DI TRENTO  
2009–currently

Teacher of *Statistical Relational Learning* (3 CFU), Doctoral Course in Information and Communication Technology

UNIVERSITA' DEGLI STUDI DI TRENTO  
2009–currently

Teacher of *Bioinformatics*, Doctoral Course in Biomolecular Sciences

2008–currently

Teacher of *Informativa* (9 CFU), Bachelor's Degree in Biomolecular Sciences and Technology.

UNIVERSITA' DEGLI STUDI DI TRENTO  
2008–currently

Teacher of *Machine Learning* (6 CFU), Master's Degree in Computer Science.

UNIVERSITA' DEGLI STUDI DI TRENTO  
2008-2009

Teacher of *Complex Systems* (3 CFU), Doctoral Course in Information and Communication Technology

UNIVERSITA' DEGLI STUDI DI FIRENZE  
2006-2008

Teacher of *Conoscenze informatiche e relazionali* (4 CFU), Bachelor's Degree in Scienze dell'Ingegneria Edile.

## Professional experiences

ITTIG - CNR ITALIA  
April 2003-April 2004

Application of machine learning techniques for classification and parsing problems in the treatment of legal documentation [15][36].

ESTMODUSINREBUS S.R.L.  
February 2002-February 2003

Realization of a medicine web portal within the project "La Medicina In Rete".

FINSYSTEM S.R.L.  
UNIVERSITA' DEGLI STUDI DI FIRENZE  
February 2001-December 2001

Technical consulting and software development for hierarchical classification of web pages by means of probabilistic classifiers.

## Main Software Implementations

- *BC-EMO* [8][45][47]: a genetic algorithm adapting to the decision maker.
- *CatANalyst* [9]: a web server for predicting catalytic residues in proteins from sequence and structure. Available as a standalone server (<http://catanalyst.disi.unitn.it>).
- *kFOIL* [39, 32][10][48]: a statistical relational learning system combining ILP and kernels.
- *TET* [30] [49]: an expressive representation language capable of encoding complex combinatorial features of relational data. With a search-based algorithm to learn TETs from data.
- *MetalDetector* [5, 12]: a web server for predicting metal binding sites and disulfide bridges in proteins from sequence. Available as a standalone server (<http://metaldetector.dsi.unifi.it>).
- *DISULFIND* [17]: a disulfide bonding state and cysteine connectivity prediction server. Available both as a standalone server (<http://disulfind.dsi.unifi.it>) and integrated into the PredictProtein service (<http://www.predictprotein.org/>).

- *Zinc Finder* [14]: a software implementing a machine learning method for the prediction of the zinc-binding state of pairs of nearby amino-acids from sequence only. Available at <http://zincfinder.dsi.unifi.it>
- *Kernels on Prolog Proof Trees* [16]: a software implementing kernels on Proof Trees generated by Prolog programs, and including a module for computing kernels over ground terms [35]. Available, with a tutorial on its application to protein classification, at <http://www.dsi.unifi.it/~paolo/ECML-Tutorial/>

### Program Co-Chair

- Machine Learning and Intelligent Optimization in Bioinformatics (Maliob) 2009, workshop at LION 2009.
- Machine Learning and Intelligent Optimization in Bioinformatics (Maliob) 2008, special session at CIBB 2008. Together with Roberto Battiti, Mauro Brunato.

### Reviewing activity

- Program/reviewing committee member:
  - ICML 2009-2011
  - AAAI 2010
  - ILP 2010
  - PRIB 2010-2011
  - IJCAI 2009,2011
  - NIPS 2008-2009
  - MLG 2007-2009
  - ECML/PKDD 2006,2007,2009-2011
- Reviewer for international journals:
  - ACM Transactions on Internet Technology
  - Artificial Intelligence Journal
  - BMC Bioinformatics
  - Electronic Letters on Computer Vision and Image Analysis
  - IEEE Transactions on Evolutionary Computation
  - IEEE Transactions on Neural Networks
  - IEEE Transactions on Pattern Analysis and Machine Intelligence
  - IEEE Transactions on Systems, Man and Cybernetics
  - IEEE Transactions on Knowledge and Data Engineering
  - Journal of Machine Learning Research
  - Machine Learning Journal
  - Neural Networks

- Neurocomputing
- Pattern Analysis and Applications
- Pattern Recognition
- Pattern Recognition Letters

## Supervised PhD students

- Elisa Cilia
- Stefano Teso
- Umut Avci

## Publications **Theses**

- [1] A. Passerini. *Kernel Methods, Multiclass Classification and Applications to Computational Molecular Biology*. PhD thesis, Dipartimento di Sistemi e Informatica, Università degli Studi di Firenze, 2004.
- [2] A. Passerini. Tecniche di apprendimento automatico applicate al recupero di informazione da internet. Master's thesis, Computer Engineering, Università degli Studi di Firenze, 2000.

## International journals

- [3] Elisa Cilia, Neils Landwehr, and Andrea Passerini. Relational feature mining with hierarchical multitask kfoil. *Fundamenta Informaticae*, 113(2):151–177, December 2011.
- [4] Andrea Passerini, Marco Lippi, and Paolo Frasconi. Predicting metal-binding sites from protein sequence. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 9:203–213, January 2012.
- [5] A. Passerini, M. Lippi, and P. Frasconi. Metaldetector v2.0: predicting the geometry of metal binding sites from protein sequence. *Nucleic Acids Res*, 39(Web Server issue):W288–92, 2011.
- [6] W. Shi, M. Punta, J. Bohon, J.M. Sauder, R. D’Mello, M. Sullivan, J. Toomey, D. Abel, M. Lippi, A. Passerini, P. Frasconi, S.K. Burley, B. Rost, and M.R. Chance. Characterization of metalloproteins by high-throughput x-ray absorption spectroscopy. *Genome Res*, 21(6):898–907, 2011.
- [7] M. Lippi, M. Jaeger, P. Frasconi, and A. Passerini. Relational information gain. *Machine Learning*, 2010.
- [8] R. Battiti and A. Passerini. Brain-computer evolutionary multi-objective optimization (bc-emo): a genetic algorithm adapting to the decision maker. *IEEE Transactions on Evolutionary Computation*, 2010.

- [9] Elisa Cilia and Andrea Passerini. Automatic prediction of catalytic residues by modeling residue structural neighborhood. *BMC Bioinformatics*, 11(1):115, 2010.
- [10] N. Landwehr, A. Passerini, L. De Raedt, and P. Frasconi. Fast learning of relational kernels. *Machine Learning*, 79(3):305–342, 2010.
- [11] F. Costa, A. Passerini, M. Lippi, and P. Frasconi. A semiparametric generative model for efficient structured-output supervised learning. *Annals of Mathematics and Artificial Intelligence*, 54(1-3):207–222, 2008.
- [12] M. Lippi, A. Passerini, M. Punta, B. Rost, and P. Frasconi. Metaldetector: a web server for predicting metal binding sites and disulfide bridges in proteins from sequence. *Bioinformatics*, 24(18):2094–2095, 2008.
- [13] M. Vincent, A. Passerini, M. Labbè, and P. Frasconi. A simplified approach to disulfide connectivity prediction from protein sequences. *BMC Bioinformatics*, 9(20), 2008.
- [14] A. Passerini, C. Andreini, S. Menchetti, A. Rosato, and P. Frasconi. Predicting zinc binding at the proteome level. *BMC Bioinformatics*, 8(39), 2007.
- [15] E. Francesconi and A. Passerini. Automatic classification of provisions in legislative texts. *Artificial Intelligence and Law*, 15(1):1–17, 2007.
- [16] A. Passerini, P. Frasconi, and L. De Raedt. Kernels on prolog proof trees: Statistical learning in the ilp setting. *Journal of Machine Learning Research (Special Topic on Inductive Programming)*, 7:307–342, 2006.
- [17] A. Ceroni, A. Passerini, A. Vullo, and P. Frasconi. Disulfind: a disulfide bonding state and cysteine connectivity prediction server. *Nucleic Acids Research*, 34(Web Server Issue):W177–W181, 2006.
- [18] A. Passerini, M. Punta, A. Ceroni, B. Rost, and P. Frasconi. Identifying cysteines and histidines in transition-metal-binding sites using support vector machines and neural networks. *PROTEINS: Structure, Functions and Bioinformatics*, 65(2):305–316, 2006.
- [19] A. Passerini, M. Pontil, and P. Frasconi. New results on error correcting output codes of kernel machines. *IEEE Transactions on Neural Networks*, 15(1):45–54, 2004.
- [20] A. Passerini and P. Frasconi. Learning to discriminate between ligand-bound and disulfide-bound cysteines. *Protein Engineering, Design and Selection*, 17(4):367–373, 2004.
- [21] A. Ceroni, P. Frasconi, A. Passerini, and A. Vullo. Predicting the disulfide bonding state of cysteines with combinations of kernel machines. *Journal of VLSI Signal Processing*, 35(3):287–295, 2003.

## Book chapters

- [22] A. Passerini. Kernel methods for structured data. In M. Bianchini, L. Jain, and M. Maggini, editors, *Handbook on Neural Information Processing*. Springer, 2012. (in press).

- [23] P. Frasconi and A. Passerini. Learning with kernels and logical representations. In L. De Raedt, P. Frasconi, K. Kersting, and S. Muggleton, editors, *Probabilistic Inductive Logic Programming: Theory and Application*, volume LNAI 4911, pages 56–91. Springer, 2008.
- [24] A. Passerini and A. Vullo. Machine learning in structural genomics. In *Bioinformatica: sfide e prospettive*. Franco Angeli Press, 2007.

## International conferences

- [25] Paolo Campigotto, Andrea Passerini, and Roberto Battiti. Active learning of combinatorial features for interactive optimization. In *LION*, pages 336–350, 2011.
- [26] S. Teso, C. Di Risio, A. Passerini, and R. Battiti. An on/off lattice approach to protein structure prediction from contact maps. In *Proceedings of Pattern Recognition in Bioinformatics (PRIB2010)*, Lecture Notes in Bioinformatics (LNBI). Springer, 2010.
- [27] F. Mascia, E. Cilia, M. Brunato, and A. Passerini. Predicting structural and functional sites in proteins by searching for maximum-weight cliques. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10)*, 2010.
- [28] M. Lippi, M. Jaeger, P. Frasconi, and A. Passerini. Relational information gain. In *19th International Conference on Inductive Logic Programming (ILP'09)*, 2009.
- [29] P. Frasconi and A. Passerini. Predicting the geometry of metal binding sites from protein sequence. In *Twenty-Second Annual Conference on Neural Information Processing Systems (NIPS'08)*, pages 465–472, 2009.
- [30] P. Frasconi, M. Jaeger, and A. Passerini. Feature discovery with type extension trees. In *18th International Conference on Inductive Logic Programming (ILP'08)*, 2008.
- [31] A. Vullo, A. Passerini, P. Frasconi, F. Costa, and G. Pollastri. On the convergence of protein structure and dynamics. statistical learning studies of pseudo folding pathways. In *6th European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics (EVOBIO'08)*, 2008.
- [32] N. Landwehr, A. Passerini, L. De Raedt, and P. Frasconi. kfoil: Learning simple relational kernels. In *Proceedings of AAAI'06*, Boston, Massachusetts, USA, July 16–20, 2006.
- [33] S. Menchetti, A. Passerini, P. Frasconi, C. Andreini, and A. Rosato. Improving prediction of zinc binding sites by modeling the linkage between residues close in sequence. In *Proceedings of RECOMB'06*, pages 309–320, Venice, Italy, April 2-5, 2006.
- [34] A. Passerini, P. Frasconi, and L. De Raedt. Kernels for logic proof trees. In *Dagstuhl Seminar 05051: Probabilistic, Logical and Relational Learning Towards a Synthesis*, 2005. (invited).

- [35] A. Passerini and P. Frasconi. Kernels on prolog ground terms. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence*, pages 1626–1627, Edinburgh, Scotland, UK, 2005.
- [36] C. Biagioli, E. Francesconi, A. Passerini, S. Montemagni, and C. Soria. Automatic semantics extraction in law documents. In *Proceedings of ICAIL'05*, pages 133–140, Bologna, Italy, June 6-11, 2005.
- [37] A. Passerini, M. Pontil, and P. Frasconi. From margins to probabilities in multiclass learning problems. In F. van Harmelen, editor, *Proc. 15th European Conf. on Artificial Intelligence*, 2002.
- [38] P. Frasconi, A. Passerini, and A. Vullo. A two-stage SVM architecture for predicting the disulfide bonding state of cysteines. In *Proc. of the IEEE Workshop on Neural Networks for Signal Processing*, 2002.
- [39] Elisa Cilia, Neils Landwehr, and Andrea Passerini. Relational feature mining with hierarchical multitask kfoil. *Fundamenta Informaticae*, 113(2):151–177, December 2011.

## National conferences

- [40] A. Ceroni, P. Frasconi, A. Passerini, and A. Vullo. A combination of support vector machines and bidirectional recurrent neural networks for protein secondary structure prediction. In A. Cappelli and F. Turini, editors, *AI\*IA 2003: Advances in Artificial Intelligence*, pages 142–153, 2003.
- [41] A. Passerini, M. Pontil, and P. Frasconi. On tuning hyper-parameters of multiclass margin classifiers. In *AI\*IA Workshop su Apprendimento Automatico: Metodi e Applicazioni*, 2002.
- [42] A. Ceroni, P. Frasconi, A. Passerini, and A. Vullo. Predicting the disulfide bonding state of cysteines with combinations of kernel machine. In *Primo Workshop Nazionale sulla Bioinformatica dell'AI\*IA*.
- [43] A. Passerini, P. Frasconi, and G. Soda. Evaluation methods for focused crawling. In *Atti del 7 Congresso dell'Associazione Italiana di Intelligenza Artificiale (AI\*IA)*, Bari, Italia, 25-28 Settembre 2001.

## Workshops

- [44] E. Cilia and A. Passerini. Frankenstein junior: a relational learning approach toward protein engineering. In *ECCB 2010 Workshop on Annotation, Interpretation, and Management of Mutations (AIMM 2010)*, Ghent (Belgium), 2010.
- [45] P. Campigotto, A. Passerini, and R. Battiti. Handling concept drift in preference learning for interactive decision making. In *ECML/PKDD 2010 Workshop on Handling Concept Drift in Adaptive Information Systems (HaCDAIS 2010)*, Barcelona (Spain), 2010.

- [46] C. Nicolini, B. Lepri, S. Teso, and A. Passerini. From on-going to complete activity recognition exploiting related activities. In *International Workshop on Human Behaviour Understanding (HBU'10)*, 2010.
- [47] P. Campigotto and A. Passerini. Adapting to a realistic decision maker: experiments towards a reactive multi-objective optimizer. In *LION workshop on Multiobjective Metaheuristics (LION-MOME)*, 2010.
- [48] Elisa Cilia, Niels Landwehr, and Andrea Passerini. Mining drug resistance relational features with hierarchical multitask kfoil. In *Proceedings of BioLogical@AI\*IA2009*, December 2009.
- [49] P. Frasconi, M. Jaeger, and A. Passerini. Learning type extension trees for metal bonding state prediction. In *ECML'08 Workshop on Statistical and Relational Learning in Bioinformatics*, 2008.
- [50] A. Passerini and P. Frasconi. Proof tree kernels: a candidate ingredient for intelligent optimization. In *Learning and Intelligent OptimizatioN - LION 2007 II*, 2007.
- [51] F. Costa, A. Passerini, and P. Frasconi. Learning structured outputs via kernel dependency estimation and stochastic grammars. In *ECML'06 Workshop on Mining and Learning with Graphs (MLG 2006)*, 2006.
- [52] F. Costa, S. Menchetti, A. Ceroni, A. Passerini, and P. Frasconi. Decomposition kernels for natural language processing. In *EACL'06 Workshop on Learning Structured Information in Natural Language Applications*, 2006.
- [53] A. Passerini, P. Frasconi, and L. De Raedt. Kernels on prolog proof trees: Statistical learning in the ilp setting. In *ICML '05 Workshop on Approaches and Applications of Inductive Programming*, 2005.
- [54] P. Frasconi, A. Passerini, S. Muggleton, and H. Lodhi. Declarative kernels. In S. Kramer and eds. B. Pfahringer, editors, *Late-Breaking Papers of the 15th International Conference on inductive Logic Programming (ILP 05)*, Bonn, Germany, 2005.

Trento, 16.01.2012