

CURRICULUM VITAE

Luís António Diniz Fernandes de Morais Sarmento

Born: 1976, March 19

Address: Rua das Andresas 88, 3.5, 4100 – 050 Porto, Portugal

Web-page: <http://www.fe.up.pt/~las/>

E-mail: las@fe.up.pt

A - Current Positions

Phd. Student in Computer Engineering - Faculdade de Engenharia da Universidade do Porto (Portugal). Phd. scholarship grant from Fundação para a Ciência e Tecnologia (Portugal) SFRH/BD/23590/2005.

Related Publications:

- Luís Sarmento, Jorge Filipe Teixeira, Eugénio Oliveira. Assessing the Impact of Thesaurus Based Expansion Techniques in QA-centric IR. In Carol Peters, Tomas Deselaers, Nicola Ferro, Julio Gonzalo, Gareth J.F.Jones, Mikko Kurimo, Thomas Mandl, Anselmo Peñas & Viviane Petras (eds.), Evaluating Systems for Multilingual and Multimodal Information Access 9th Workshop of the Cross-Language Evaluation Forum, CLEF 2008, Aarhus, Denmark, September 17-19, 2008, Revised Selected Papers 2008, Springer.
- Luis Sarmento, Valentin Jijkuon, Maarten de Rijke, Eugénio Oliveira. "More like these": growing entity classes from seeds. Proceedings of the sixteenth ACM Conference on Information and Knowledge Management (CIKM '07). Lisbon, November 2007, ISBN:978-1-59593-803-9.
- Luís Sarmento "A first step to address biography generation as an iterative QA task". In Carol Peters, Paul Clough, Fredric C. Gey, Jussi Karlgren, Bernardo Magnini, Douglas W. Oard, Maarten de Rijke & Maximilian Stempfhuber (eds.), Evaluation of Multilingual and Multi-modal Information Retrieval - 7th Workshop of the Cross-Language Evaluation Forum, CLEF 2006. Alicante, Spain, September, 2006. Revised Selected papers. Berlin / Heidelberg: Springer, Lecture Notes in Computer Science, 2007, pp. 473-482.
- Luís Sarmento. "BACO - A large database of text and co-occurrences". In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'2006). Génova, Italia, 22-28 May 2006
- Luís Costa & Luís Sarmento. "Component Evaluation in a Question Answering System". In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'2006). Génova, Italia, 22-28 May 2006
- Luís Sarmento. "SIEMÊS - a Named-Entity Recognizer for Portuguese Relying on Similarity Rules ". In Renata Vieira, Paulo Quaresma, Maria da Graça Volpes Nunes, Nuno Mamede, Claudia Oliveira & Maria Carmelita Dias (eds.), Encontro para o Processamento Computacional da Língua Portuguesa Escrita e Falada (PROPOR'2006). Itatiaia, RJ, 13 -17 May
- Luís Sarmento, Ana Sofia Pinto & Luís Cabral. "REPENTINO - A collaborative

wide-scope gazetteer for Entity Recognition in Portuguese". In Renata Vieira, Paulo Quaresma, Maria da Graça Volpes Nunes, Nuno Mamede, Claudia Oliveira & Maria Carmelita Dias (eds.), Encontro para o Processamento Computacional da Língua Portuguesa Escrita e Falada (PROPOR'2006). Itatiaia, RJ, 13 - 17 May.

B - Past Positions

Google Inc. (www.google.com)

Engineering Intern at NY offices. Main task:

- Developing and experimenting Large Scale Named-Entity Disambiguation techniques, using Google's Map-Reduce framework

Linguateca (www.linguateca.pt)

Researcher at Linguateca from October 2002 to December 2005. Main tasks:

- Development of Corpógrafo, a web-based environment for corpora research and knowledge engineering.
- Development of EVAL, METRA, web tools for studying Machine Translation and for compiling a corpus of human-evaluated machine translation sentences.
- Development of SIEMÊS, a named-entity recognition system for Portuguese

Related Publications:

- Luís Sarmiento, Belinda Maia, Diana Santos, Ana Pinto & Luís Cabral "Corpógrafo V3: From Terminological Aid to Semi-automatic Knowledge Engine". In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'2006) Génova, Italia, 22-28 May 2006
- Débora Oliveira, Luís Sarmiento, Belinda Maia & Diana Santos. "Corpus analysis for indexing: when corpus-based terminology makes a difference". In Pernilla Danielsson & Martijn Wagenmakers (eds.), Proceedings of the Corpus Linguistics 2005 Conference Birmingham, UK, 14-17 July 2005. ISSN: 1747-9398
- Luís Sarmiento. "A Simple and Robust Algorithm for Extracting Terminology". In META Symposium - For a Proactive Translatology Université de Montréal, Québec, Canadá, 7-9 de Abril de 2005
- Belinda Maia & Luís Sarmiento. "The Corpógrafo - an Experiment in Designing a Research and Study Environment for Comparable Corpora Compilation and Terminology Extraction". In Proceedings of eCoLoRe / MeLLANGE Workshop, Resources and Tools for e-Learning in Translation and Localisation Centre for Translation Studies, University of Leeds, UK, 21-23 March 2005
- Belinda Maia, Luís Sarmiento & Diana Santos "Introduzindo o Corpógrafo - um conjunto de ferramentas para criar corpora especializados e comparáveis e bases de dados teminológicas". Terminómetro - Número especial nº 7 - A terminologia em Portugal e nos países de língua portuguesa em África (2005): 61-62.
- Diana Santos et al. "Linguateca: um centro de recursos distribuído para o processamento computacional da língua portuguesa". In Guillermo De Ita Luna, Olac Fuentes Chávez and Mauricio Osorio Galindo (eds.), Proceedings of the

international workshop "Taller de Herramientas y Recursos Lingüísticos para el Español y el Portugués", IX Iberoamerican Conference on Artificial Intelligence (IBERAMIA 2004), Puebla, México, 23 November 2004, pp. 147-154.

- Luís Sarmento, Belinda Maia and Diana Santos. "The Corpógrafo - a Web-based environment for corpora research", Proceedings of the Language and Resources Evaluation Conference, 2004, Lisbon, Portugal, 25 May 2004, pp. 449-452.
- Diana Santos, Belinda Maia and Luís Sarmento. "Gathering empirical data to evaluate MT from English to Portuguese", Proceedings of the Workshop on the Amazing Utility of Parallel and Comparable Corpora, held at the Language and Resources Evaluation Conference 2004, Lisbon, Portugal, 25 May 2004, pp. 14-17.
- Belinda Maia and Luis Sarmento, "Gestor de Corpora – Um Ambiente Web integrado para Linguística baseada em Corpora", Almeida J.J. (Ed.) Actas da CP3A: Corpora Paralelos, Aplicações e Algoritmos Associados, Braga, Junho de 2003.
- Diana Santos and Luís Sarmento. "O projecto AC/DC: acesso a corpora/disponibilização de corpora". In Amália Mendes and Tiago Freitas (orgs.), Actas do XVIII Encontro da Associação Portuguesa de Linguística (APL 2002) (Porto, 2-4 Outubro 2002), Lisboa: APL, pp. 705-717.

Faculdade de Engenharia de Universidade do Porto (www.fe.up.pt)

Assistant lecturer at the Faculty of Engineering of the University of Porto (FEUP), from October 2003 to October 2005. Teaching computer science disciplines at the undergraduate and MSc. levels: Algorithms and Data Structures, Database Programming and Distributed AI.

Escola das Artes - Universidade Católica Portuguesa (www.artes.ucp.pt)

Assistant Lecturer at the Art School of the Catholic University in Porto, from October 1999 to August 2002, teaching Multimedia and Web programming.

INESC (www.inescporto.pt)

Internship in the Video Coding Group at INESC (Instituto Nacional de Engenharia e Sistemas de Computadores), from October 1997 to September 1998. Main task: development of encryption tools for MPEG2 video streams.

Related Publications:

- Luis Teixeira and Luís Sarmento. "Secure transmission of MPEG video sources", Proceedings [ISPACS'98](#) - 6th IEEE International Workshop on Intelligent Signal Processing & Communications System - Melbourne, Australia, 04-06 November 1998
- Luis Sarmento, Luis Teixeira and Artur P. Alves. "Guaranteeing Confidentiality and Integrity of compressed video sources". Proceedings of [RECPAD'98](#) - 10th Portuguese Conference on Pattern Recognition - IST, Lisboa, Portugal, 26-27 March 1998.

C - Academic Background

FEUP – Faculdade de Engenharia da Universidade do Porto (www.fe.up.pt):

Graduated in Electrical and Computer Engineering in July 1999.

Final Grade: 16/20 - Top 5%.

FCUP – Faculdade de Ciências de Universidade do Porto (www.fc.up.pt):

MSc. in Artificial Intelligence. Dissertation theme: Emotion-based Agent Architectures.

Final Grade: A+.

Related publications:

- David Pereira, Eugénio Oliveira, Nelma Moreira and Luís Sarmiento "Towards an Architecture for Emotional BDI Agents". In IEEE Proceedings of EPIA, Covilhã, December 2005.
- Luís Sarmiento, Daniel Moura and Eugénio Oliveira. "Fighting Fire with Fear", in Proceeding of the 2nd European Workshop on Multi-Agent Systems (EUMAS 2004), Barcelona, 16-17 December 2004.
- Luís Sarmiento. "An Emotion-based Agent Architecture". Msc. Thesis. Faculdade de Ciências da Universidade do Porto. Portugal, May 2004
- Eugénio Oliveira and Luís Sarmiento. "Emotional advantage for adaptability and autonomy". In Proceedings of the second international joint conference on Autonomous Agents and Multi-Agent Systems (AAMAS'03), Melbourne. pages 305–312, New York, USA, 2003. ACM Press.
- Eugénio Oliveira and Luís Sarmiento. "Emotional Valence-based Mechanisms and Agent Personality", in Proceedings of SBIA'02 - Simpósio Brasileiro de Inteligência Artificial, Recife, Brasil, November 2002. Lecture Notes in Artificial Intelligence, Springer Verlag.