

Cristiano Castelfranchi

CURRICULUM

1. Some general data

Cristiano Castelfranchi is born in Roma- Italy l'8 giugno 1944

Laurea in Humanities at the University of Roma "La Sapienza" nel 1969, Faculty of Humanities and Philosophy, with a dissertation on "The representation of meaning" (110/110 e lode) Prof. Rodolfo Nencini (Psychology), and Prof. Tullio De Mauro (Philosophy of Language).

Working since 1965 at the Institute of Psychology of the Italian National Research Council, with Prof. Domenico Parisi.

Head of the Division of "AI, Cognitive and Interaction Modeling", and of the Division of "Social Psychology" of the Institute of Psychology of CNR – Roma

Since 2001, full professor of "Cognitive Sciences" (General Psychology) at the University of Siena, Department of Communication Science; and

since 2002, Director of the Institute of Cognitive Sciences and Technologies of the National Research Council (ISTC-CNR) in Rome.

2. Books

14 books with Italian publishers. Among them:

- "Linguaggio, conoscenze e scopi" (*Language, knowledge and goals*) (with D. Parisi). Bologna: Il Mulino, 1980;
- "Il linguaggio e la macchina" (*The Language and the machine*) (with D. Parisi). Torino: Boringhieri, 1987;
- "Che figura: Emozioni e Immagine Sociale" (*Emotions and Social Image*), Bologna: Il Mulino, 1988; (seconda edizione, 2005)
- "La cognizione del valore" (*The Cognition of Value*) (with M. Miceli), Milano: Franco Angeli, 1992;
- "Sensi di colpa" (*Guilt Feelings*) (con I. Poggi e R. D'Amico), Giunti, Firenze, 1995;
- "Le difese della mente" (*The Cognitive Psychology of Defensive Mechanisms*) (con M. Miceli), La Nuova Italia Scientifica, Roma, 1995;
- Per una Psicologia della Riabilitazione nella Crisi della Psichiatria Istituzionale (con P. Henry e A. Pirella) Edizioni Gruppo Abele, Torino, 1995
- "La società delle menti" (*Society of Minds*) (with R. Conte), UTET, Torino, 1996.
- "Bugie, finzioni e sotterfugi. Per una scienza dell'inganno" (*For a Science of Deception*) (con I. Poggi). Roma, Carocci, 1998 (seconda edizione, 2000).
- "Fondamenti del Cognitivismo Clinico" (con M. Miceli e F. Mancini). Bollati-Boringhieri. Torino, 2004

6 books with international publishers:

- “Cognitive and Social Action” (co-author R. Conte), UCL Press (Taylor& Francis), London, 1995.
- “Artificial Social Systems”, (co-editor E. Werner), Springer, LNAI, Berlin, 1994;
- “From Reaction to Cognition” (co-editor J.P.Muller), Springer, LNAI, Berlin, 1995;
- “Trust and Deception in Virtual Societies” (co-editor Y-H. Tan), Kluwer, Dordrecht, 2002
- “Intelligent Agents VII: Theories, Architectures, Languages” (co-editor Y. Lesperance) - Springer, LNAI, Berlin 2002
 - Autonomous Agents and Multi-Agent Systems '09 (con C. Sierra e J. Sichman) Budapest, 2009
 - “Trust Theory” (co-author R. Falcone) John Wiley & Sons, UK, 2010
- Expectancy & Emotions (con M.Miceli) Oxford University press, 2015

More than 250 papers in journals, books, proceedings of international conferences. (see below).

3. Responsibility of Research Projects (international)

Bilateral Projects (international cooperation):

"Man-Machine Interface" USA (1987-88); “Dependency and Psychological Wellbeing” Prof. S. Brehm (university of Kansas, Lawrence, USA) (1992-93); "Intelor" (Intelligent Organisations) with Portugal-INESC (Prof. H. Coelho); "Cognitive Modeling" with the Bulgarian Academy of Science - BAS, Sofia, (dr. Kokinov) (1995-97 e 1998-2000);

European Projects:

- Project "ModelAge"-ESPRIT (1994-97) - Unit;
 - Unit of ICSA European project (in collaboration with the Univ. of Siena, resp. UMIST- Manchester- Prof. O'Hare) Information Cooperative for Satellite in CEO- Consortium for Earth Observation, 1994-96;
 - Unit of the European Network of Excellence *AgentLink* (1998-99);
 - Unit in Project ALFEBIITE (A Logical Framework for Ethical Behaviour between Infohabitants in the Information Trading Economy of the Universal Information Ecosystem), V Framework, (October 1999- November 2000).
- Coordinator EU Strep Project, Contract number: IST-511931; Funded under 6th FWP (Sixth Framework Programme; "From Reactive to Anticipatory Cognitive Embodied Systems" (MindRACES) (2004-2007), *first phase*.
 - Unit in the ESF (European Science Foundation) The Origin of Man, Language and Languages (OMLL) (EUROCORES Programme OMLL, Contract no. ERAS-CT-2003-980409).
 - Unit HUMAINE (Human-Machine Interaction Network on Emotion) - 2004/2007 - Network of Excellence in the EU's Sixth Framework Programme [IST Thematic Priority IST-2002-2.3.1.6 Multimodal Interfaces]
 - Unit in the ESF (European Science Foundation) CNCC on “Consciousness” [IST Thematic Priority].
 - Coordinator of the proposal "The Social and Mental Dynamics of Cooperation" for the ESF call "The Evolution of Cooperation and Trade", with Samuel Bowles, Ernst Fehr, Herbert Gintis, Michel Kerszberg, Ruth Mace, Ugo Pagano, and others (waiting for grants)

CONTACT - Consciousness in Interaction: The Role of the Natural and Social Environment in Shaping Consciousness

CONTACT is a Collaborative Research Project (CRP) running from 2007 to 2009 under the ESF Programme *Consciousness in a Natural and Cultural Context*. The project is co-funded by the European Science Foundator

and by three national funding agencies: the CNR (Italy), the AHRC (United Kingdom), and the NWO (The Netherlands).

Our research is built on the assumption that the brains and bodies of cognitive agents (humans and animals) interact dynamically with both their natural and social environments. The CONTACT project questions the assumption that conscious experience must be explained by the brain by itself, as opposed to the embodied brain in interaction with environments, both natural and social.)

Research Units

Project Leader

Prof. Cristiano Castelfranchi, Istituto di Scienze e Tecnologie della Cognizione - CNR, Roma (IT)

Principal Investigators *Dr. Finn Spicer⁽¹⁾, University of Bristol (UK) Prof. Andy Clark, University of Edinburgh (UK) Prof. Ed S. Tan, Universiteit van Amsterdam (NL) Prof. Enrico Rambaldi, Istituto per la Storia del Pensiero Filosofico e Scientifico Moderno - CNR, Milano (IT)*

Associate Partner *Prof. Thomas Metzinger, Johannes Gutenberg Universität Mainz (DE)*

4. INVITED SPEAKER

4.1. 'Invited speaker' or 'keynote speaker' in international conferences or meetings

- *IJCAI - 97 International Joint Conference of Artificial Intelligence: Nagoya, Japan, 1997 - "Modeling Social Action for AI Agents"*
 - MAAMAW'97 European Workshop on Modeling Autonomous Agents in a Multi-Agent World; Roneby, maggio 1997 invited talk on "Limits of Economic and Strategic Rationality for Agents and Multi-Agent Systems"
 - SimSoc'97 Workshop on Simulating Societies, Cortona, 20 sett 97, invited talk on "Challenges for agent-based social simulation: social functions"
 - Fourth JURAD Meeting - European project; Madeira 28-29 March 1994: invited talk su "Social Commitments"
 - *Journées Francophones sur l'Intelligence Artificielle Distribuée et les Systèmes Multi-Agents*, Chambéry, marzo 1995 ("Dependence Theory")
 - III FOMAS workshop *Foundations of Multi-Agent Systems*, Warwick dic 97 "Cognitive Social Agents and MAS"
 - International Workshop on "Learning in DAI Systems" at the European Conference on AI - ECAI 96
 - First European Agent Systems Summer School - EASSS'99, Amsterdam, July 1999
 - Workshop "AOIS'99- Agent oriented information systems", Heidelberg, June 1999 invited talk on "Trust and Deception in Artificial Societies"
 - IBERAMIA'98 - 6th Ibero-American Conference on Artificial Intelligence, Lisbon, October 1998;
 - Workshop on "Collective Intentionality", Monaco June 1999, Institute of Logic and Philosophy. Invited talk on "Cooperation and Commitment"
 - Workshop "Emotions in Humans and Artifacts", Vienna, August 1999
 - International Workshop on "Affect in Interactions: Towards a New Generation of Interfaces", Siena, October 1999
 - Workshop "Strategies for Large Scale Emergent Computing Systems", UK, Cardiff, July 2000
 - Symposium on "Argumentation and Computation", UK, June 2000
 - Workshop su Agent technology, Taiwan, October 2000
 - Workshop "Foundations and Applications of Collective Agent Based Systems" at ESSLLI-99 European Summer School on Logics and Language, Utrecht, August 1999
- Workshop "DEON'00" Deontic Logic in Computer Science, Toulouse January 2000
- Workshop "2° SARA Workshop" - Lisbon 11-2- 2000 "Power, Delegation, Responsibility"
 - Workshop "Agents' conflicts"-AAAI99, USA, Orlando, July 1999
- 6th International Summer School in Cognitive Science- Sofia, 12 - 31 July 1999

- <<http://www.nbu.bg/cogs/events/ss99.html>><http://www.nbu.bg/cogs/events/ss99.html>
- Workshop “Socially Intelligent Agents - The Human in the Loop” - AAAI-FS, November 2000 Boston
 - International Joint Conference IBERAMIA'2000-SBIA'2000 (Ibero-American Conference on Artificial Intelligence & Brazilian conference on AI) - November 19-22, 2000 Sao Paulo, Brazil
 - 7th Scandinavian AI conference- Odense, 19-21 February 2001
 - "CIA - Cooperative Information Agents", Modena- Settembre 20001
 - Complexity in Social Science - COSI Summer School 2002- 30 Giugno/ 6 Luglio 2002, Chania, Creta. <<http://www.irit.fr/COSI/summerschool/index.html>><http://www.irit.fr/COSI/summerschool/index.html>
 - *Summer School in Economics “Cognitive Processes and Economic Behaviour”* (organizers Nicola Dimitri., Marcello Basili & Itzhak Gilboa)- Siena – Pontignano- 2002
 - Symposium on “Cognition and Rationality”; Padua, 17-21 March, 2003 Invited talk on “Rational Decisions, Human Motives, and Emotions”.
 - Convegno AIFA- “Mental Processes, Communication and Rules”, Vietri, 2-3 Ottobre 2003 <<http://sifa.unige.it/vietri/index.html>><http://sifa.unige.it/vietri/index.html>
 - Symposium on the Foundation of Interaction Design, Ivrea, 12-13 Novembre 2003 <<http://www.interaction-ivrea.it/en/news/education/2003-04/symposium/index.asp>><http://www.interaction-ivrea.it/en/news/education/2003-04/symposium/index.asp>
 - 1st International Workshop on Socio-Cognitive Grids (<http://www soi.city.ac.uk/~kostas/sc-grids/cfp.html>). In: Tales of the Disappearing Computer convention, 1-4 of June 2003, Santorini, Greece
 - SIKS day 2004 – Utrecht; invited talk on “Artificial Social Intelligence”
 - workshop "Changing minds: Cognitive, computational and logical approaches to belief change", University of Amsterdam, 29 Ottobre 2004, Netherlands Organisation for Scientific Research (NWO)
 - International WS LEA'05 “Law and Electronic Agents”- Edinburgh, 2005;
 - Tenth International Conference on ARTIFICIAL INTELLIGENCE and LAW (ICAAIL 2005), 6-11 Giugno 2005, Bologna <http://www.wogli.unibo.it/icail05/?page_name=invspk>http://www.wogli.unibo.it/icail05/?page_name=invspk
 - “Artificial Economics”- Lille, France, September 15-16 2005
 - “BVAI- Brain Vision and Artificial Intelligence” 19-21 October, Naples, Italy
 - Workshop "Minds in Interaction", Dutch Institute for High Studies (NIAS), Wassenaar, Olanda, 26-27 Settembre 2006
 - Fifth International Conference on Collective Intentionality (Collint V)- Helsinki, 2006
 - European Conference on Cognitive Science- CogSci; Delphi, Grecia, maggio 2007
 - ACII- International Conference on Affective Computing and Intelligent Interaction – Lisbon – September 2007
 - Expectation Workshop - Science and Engineering Research Council (SERC) Agency of Science, Technology and Research (A*STAR) – Singapore, Aug 21-24 2007
 - International Ws on Theory of Mind, Marina del Rey, Institute for Creative Technology, CA. October 2007
 - Advanced school *Minds & Societies*, Montreal, with D. Dennett, L. Sanger, D. Roy, J. Byrne, 2008
 - Advanced school *Social Norms*, San Sebastian, with J. Elster, D. Sperber, H. Gintis, C. Bicchieri, 2008
 - Advanced school *Social Cognition and Social Narrative*, San Marino, with S. Gallagher, S. Stich, 2008
 - Brazilian International Meeting on *Cognitive Science* (EBICC), Campinas, with M. Dascal, C. Sinha, 2009
 - IWCogSc-10 International WS on Cognitive Science, San Sebastian, ILCLI (UPV-EHU), 2010
 - 3rd International Conf. on *Agents and Artificial Intelligence* (ICAART), Rome, **2011**
- <http://vimeo.com/23900594>

- **EPIA 2011** Portuguese Conference on AI. Lisbon October 10-13 2011
- EUMAS'16 Valencia, dicembre 2016 EUMAS & AT conference

4.2 Panels (published)

Workshop ATAL'96 "Agents Theories Architectures and Languages" panel on "What is an agent" In J. Muller, M. Wooldridge, N. Jennings (Eds.) *Intelligent Agents III*, Springer, Berlin, 37-41

ICSS&SS International Conference on Social Simulation and Social Sciences, Cortona 21-27 sett 1997, panel on "Cooperation" (partially published in *Journal of Artificial Societies and Social Simulation*, 1(1), 1998)

III FOMAS workshop *Foundations of Multi-Agent Systems*, Warwick Dic 97, panel on "Social Rationality" (with K. Binmore, M. Wooldridge, ..) published in *Knowledge Engineering Review* 13, 3, 1998, 309 314

5. CHAIR or ORGANIZER of WORKSHOPS, CONFERENCES, SYMPOSIA, ...; BOARDS; REVIEWERS; etc.

5.1. 'General' 'scientific' o 'program' chair of international events

- First International Joint Conference AA-MAS, (Autonomous Agents and Multi-Agent Systems) 2002 - Program chair (with Lewis Johnson)
- International Workshop on "Agent Theories, Architectures, and Languages"- ATAL 2000, Boston.(co-chair Y. Lesperance)
- MAAMAW'92- European Ws on Modeling Autonomous Agents in a Multi-Agent World, Viterbo, August 1992 (general co-chair, con E. Werner)
- MAAMAW'93- European Ws on Modelling Autonomous Agents in a Multi-Agent World, Neuchatel, August 1993 (general co-chair, con J.P. Muller)
 - Atelier "Cognition Collective" (co-chair con Ch. Lenay) per la first European Conference on Cognitive Science, 1994
- First Interdisciplinary Workshop on "Deception, Fraud and Trust in Agent Societies" Minneapolis/St Paul, USA, May 9, 1998 - associated with AUTONOMOUS AGENTS'98 (co-chair)
- Second Interdisciplinary Workshop on "Deception, Fraud and Trust in Agent Societies" May - 1999 - associated with The Third International Conference on AUTONOMOUS AGENTS (Agents '99) Seattle (WA) - USA (co-chair)
- Interdisciplinary Workshop on "Context" at the *European Conference on Cognitive Science* - ECCS'97, Manchester UK (organizzatore e co-chair con B. Kokinov, F. Giunchiglia, P. Bouquet)
- 2° International Workshop "MASHO - MODELLING ARTIFICIAL SOCIETIES and HYBRID ORGANIZATIONS" 2002 (co-chair, with Jonker and Linderman)
- Chair of the ESF Exploratory Workshop "Understanding the Dynamics of Knowledge" 17-19 Novembre 2005, Certosa di Pontignano, Siena (IT) - <http://linus.media.unisi.it/cirg/udk/>
- Chair of the AAI 2005 Fall Symposium "From Reactive to Anticipatory Cognitive Embodied Systems" 3-6 Novembre 2005, Hyatt Crystal City in Arlington, Virginia (USA)- <http://www.mindraces.org/events/fss05/>

- Co-Chair of the AAMAS 2005 Workshop "Trust in Agent Societies" 25 Luglio 2005, Utrecht (NL)
 - <http://www.istc.cnr.it/news/trust2005.htm>
- co-Chair of the Fourth International Conference on Collective Intentionality “Cognition and Collective Intentionality Mental Prerequisites and Attitudes for Collective Intentionality: Philosophical, Logical, Computational, Neural and Psychological Models” (Collint IV)- Siena, 13/15 Ottobre 2004-
<http://www.istc.cnr.it/collintIV/committee.htm>
- General co-chair of AAMAS 2009 (with Carles Sierra) – Budapest: International Conference on Autonomous Agents and Multi-Agent System
- Chair of the OMLL International Workshop "Evolving communication: from action to language. An 'implicit vs explicit' cognitive and pragmatic perspective"- 20-21 Maggio 2004, Graduate College S. Chiara, Siena (IT)
http://www.unisi.it/ricerca/dip/fil_sc_soc/dot-sc/evcoimes.html

-
- Area-chair for Autonomous Agents and MAS; European Conference on Artificial Intelligence - ECAI 94 - Vienna;
 - Tutorials Chair per International Conference on Multi-Agent Systems - ICMAS'98, Paris
 - Chair of many sections, tracks and panels in international conferences

5.2 Organization of International Workshops, Meetings, Panels

a part those on point 5.1 (Chair)

- International Workshop on "Simulating Societies" SimSoc'93 - Certosa di Pontignano (organization cochair with N. Gilbert e J. Doran)
 - panel on "Commitment" at ICMAS'95 San Francisco;
 - panel on the notion of "Emergence" at SimSoc'93, Pontignano;
 - panel on “Cooperation and cognition”, ECCS'99, Siena.
 - panel on Anticipation, Expectations, and Surprise, at the AAAI 2005 Fall Symposium “From Reactive to Anticipatory Cognitive Embodied Systems” 3-6 Novembre 2005, Hyatt Crystal City in Arlington, Virginia (USA)- <http://www.mindraces.org/events/fss05/>

5.3 Editorial roles and boards

- Member of the CogNet Academic Council - MIT Press [14 membri tra cui Michael Gazzaniga (Dartmouth College), Kelly Jaakkola (Boston), Terrence J. Sejnowski (San Diego), Josh Tenenbaum (Stanford University)]
- Since the foundation one of the “executive editors” del “*Journal of Autonomous Agents and Multi-Agent Systems*” Kluwer Academic Publishers in Norwell, MA USA
- Since the foundation member of the ‘editorial board’ of *Cognitive Science Quarterly* (CSQ), Hermes Science Publications, Paris and Oxford
- Since the foundation member of the ‘editorial board’ of JASSS, “Journal of Artificial Societies and Social Simulation” , University of Surrey, UK.
- Since the foundation member of the ‘editorial board’ of “*Mind & Society*” (Fond. Rosselli; LUISS; Bocconi); Now: Springer.

- Guest editor (with Falcone R., Firozabadi B., Tan Y) of two Special Issues on "Trust in Agents" del *Applied Artificial Intelligence Journal*, Taylor and Francis 2000, 14 (8).
 - Guest editor (with M. Miceli) of a Special Issue on "Desires, Goals, Intentions, and Values: Computational Architectures" del journal *Cognitive Science Quarterly*, 2002
 - Guest della special issue su "Cognitive Systems" n.53, April 2003 di "ERCIM News" - dell' European Research Consortium
 - Guest editor (with L. Tummolini) of the double special Issue su "Collective Intentionality" del journal "Cognitive Systems Research", Elsevier, <http://www.cogsci.rpi.edu/~rsun/journal.html>
 - Editor (with M. Butz, R. Falcone, G. Pezzulo) of a book on Anticipatory Systems – Springer (in press)
- Scientific board member of Springer Series: Law, Governance and Technology (LGTS) Casanovas & Sartor (eds.).

5.4 Review, Evaluation, and Recommendation activity

Journals:

- *Artificial Intelligence journal*;
- *Cognitive Science Quarterly*
- JASSS, "Journal of Artificial Societies and Social Simulation"
- *Journal of Autonomous Agents and Multi-Agent Systems*
- *European Journal of Social Psychology*;
- *The Knowledge Engineering Review*;
- *Revue d'Intelligence Artificielle*;
- *International Journal of Human-Computer Studies*;
- *Artificial Intelligence and Law*;
- *Journal of Logic and Computation*
- *Computational and Mathematical Organization Theory*
- *Studies in Communication Sciences*
- *Sistemi Intelligenti (Intelligent Systems)- Il Mulino - Italia*

Conferences and Workshops:

AI*IA, EPIA, ECCS, ECAI, IJCAI, ICMAS, ATAL, MAAMAW, AUTONOMOUS AGENTS, AAMAS, ICSS&SS; SimSoc, AAI, ISATA, DEON, FOMAS, EUMAS, AGENTLINK ws, ECAI, SBIA/IBERAMIA, iTRUST, MATES; MABS; WS on Attitudes, Personalities and Emotions; CEEMAS; COOP; MASHO; CIA; LEA; Anticipatory Behavior in Adaptive Learning Systems; CONTEXT; MOO; CoOrg; NORMAS; SOAS; AISB Ws; Ambient Intelligence; and many others.

Evaluation:

International PhD theses:

- UMIST – University of Manchester - UK, 1995;
- Polytechnic of Grenoble, Francia, 1994;
- SIKT at University of Karlskrona/Roneby, Svezia
- Department of Social Psychology, University of Helsinki, Finland, 2000
- Dep. of Computer Science, University of Stockholm, Svezia, 2000
- SIKS, Dutch Research School for Information and Knowledge Systems, 2005;
- Technical University of S. Etienne, France
- University of Utrecht;
- University of Lisbon;
- New Bulgarian University; Sofia

Centers, Projects and Candidates:

- the Vrije University di Amsterdam;
- the University of Wisconsin USA;

- the INRIA Francia;
- University of Limerick;
- University College Dublin;
- Queen Mary and Westfield College, London;
- York University, Toronto, Canada
- LabMAg – Lisbon, Advisory Committee,
- and others

Recommendation letters (for awards, positions, etc.) for M. Georgeff; V. Lesser; N. Jennings; A. Ortony; C. Jonker; Y. Demazeaux; E. Werner; F. Brazier; and others

5.5 Boards of Conferences and Ws

- European Conference of Cognitive Science - ECCS 94, 96, 99;
- European Conference of Artificial Intelligence - ECAI'94, 96;
- International Conference on Multi-Agent Systems - ICMAS (since the beginning and many times)
- International Conference Autonomous Agents'01 - May-June, Montreal, 2001
- International Workshop on Agents Architectures, Theories, and Languages - ATAL dal 1995 al 2001;
- 26° Conferenza ISATA (1993)
- Conference on Cooperative Knowledge Based Systems (CKBS-94);
- International Conference on Social Simulation and Social Sciences - ICSS & SS 1997;
- Conferenza Portoghese di Intelligenza Artificiale - EPIA '97;
- Interdisciplinary Conference of Modeling and Using Context - CONTEXT 2001.
- Second Asia-Pacific Conference on Intelligent Agent Technology - IAT 2001 - Japan- October 2001
- Congresso dell'Associazione Italiana di Intelligenza Artificiale - AI*IA 89 Trento, 93 Torino, 97 Roma, Bari 01;
- European Workshop on Modelling Autonomous Agents in a Multi-Agent World - MAAMAW dal 1991 al 2001;
- di numerosissimi altri WS internazionali su “Agents”, su “Emotions”, su “Imitation”, “Conflicts”, etc. (vedi documentazione).
- Workshop ESAW'00 - Engineering Societies in the Agents' World - ECAI, European Conference of Artificial Intelligence - Berlin, August 2000
 - Workshop “Agent-Oriented Software Management” - ECAI European Conference of Artificial Intelligence - Berlin, August 2000
 - Workshop MASHO1- Modelling Artificial Societies and Hybrid Organizations - ECAI, European Conference of Artificial Intelligence - Berlin, August 2000
 - Workshop PRIMA 2000 Pacific Rim Multi-Agent Ws, Melbourne, August 2000
- Workshop “Computational Models of Natural Language Argument” - ICCS Conference 2001 - S. Francisco May 2001
 - Workshop “Norms and Institutions in MAS” Autonomous Agents Conference, Montreal, May 2001

Member of the *advisory board*

- (since the beginning) International Conference on Multi-Agent Systems;
- International Workshop on Agents Architectures, Theories, and Languages - ATAL;
- (since the beginning) European Conference of Cognitive Science - ECCS

Member of the *advisory board* of the International Summer School on Cognitive Science (Sofia - New Bulgarian University and BAS)

advisor for ICCS-99- Six International Colloquium on Cognitive Science, San Sebastian, May, 1999

advisor for **CogSci2005** - XXVII Annual Conference of the Cognitive Science Society'05 – Stresa - Italy

Founder and coordinator (for 4 years) of the Group of Interest on Agents and DAI of the *AI*IA Italian Association for Artificial Intelligence*.

Member (since the beginning) of the board of IFMAS – International Foundation for Multi-Agent Systems.

6. FORMATION ACTIVITY

LUISS University; Faculty of Economics – Roma 2007-08, 2008-09, 2009-10 ‘Psychology and Economics’

Summer Schools

- *International Summer School on Cognitive Science* (Sofia, New Bulgarian University):
 - “Emotion and Cognition” - 1995 (15 ore);
 - “Understanding Social Dependence and Cooperation” - 1999 (15 ore).
- the 9th *Autumn School on Cognition* (September 10 to September 15, 2000) at the University of Freiburg, Germany, organized by the DFG Doctoral Programs 'Human and Machine Intelligence. Corso su “Conceptual Foundations for Artificial Social Intelligence”
- European Agent Systems Summer School (EASSS 2000) (invited)
- ESSLLI-99 European Summer School on Logics and Language, Utrecht, August 1999 (invited)

Tutorials

(with Prof. M. Georgeff) tutorial on "Distributed Agent-Oriented Systems: Theory and Application" European Conference on Artificial Intelligence, Budapest, 1996

(with Rino Falcone) tutorial on “Foundations of Social Artificial Intelligence” - International Conference of Multi-Agent Systems- ICMAS, Boston, 2000.

(with Prof. F. Giunchiglia) tutorial on "I sistemi multi-agente" (MAS) Congresso Italiano di Intelligenza Artificiale- AI*IA, Roma, settembre 1997

Lesson in the seminar series on "Perspectives on Agents: Theories and Technologies" from January to June 2000 ITC-IRST/University of Trento.

ESSLLI-06 Anancy, France: Tutorial on “The Theory of Trust and MAS” (with Falcone).

PhD students (a part from Italy)

(see later)

7. Scientific Associations and Committees

Member of:

- AAAI; ACM; ECCAI; ISSA (Int. Soc. for the Study of Argumentation); IFMAS (Int. Foundation for Multi-Agent Systems)
- Committee of the Cognitive Science Society for the nomination of Fellows;
- Committee of AAMAS for Membership
- Working group of the Cognitive Science society for coordination in Europe (with Young, Vosniadou, Kokinov, etc.)

From 1989 al 1995 member of Consiglio Direttivo della associazione Italiana per l'Intelligenza Artificiale (AI*IA);

8. HONOURS and RECOGNITIONS

- PhD Honoris Causa in Cognitive Sciences, University of Turin; Cognitive Science School; 2007
 - Special award “Mente e Cervello” (Mind and Brain) for research in Cognitive Science-University of Turin; Cognitive Science School; 2007
 - Service Award of ACM - Association for Computing Machinery (per la Chairship di AAMAS’02)
 - Fellowship of ECCAI – European Coordinating Committee for Artificial Intelligence (August 2003) “for pioneering work in the field of AI”
 - Invited visitor by the Netherlands Organization for Scientific Research (NWO) at Wassenaar Center (Netherlands) for 2-4 months (NIAS Project)
 - Members ‘promoters’ of the Italian Association for Artificial Intelligence (AI*IA)
 - Honorary members of SITCC- Societa’ Italiana di Terapia Comportamentale e Cognitiva – 10/10/1991
 - Permanent member of the *advisory board* of the International Summer School on Cognitive Science (Sofia - New Bulgarian University and BAS).

-
- Evaluation received by the Institute of Cognitive Sciences and Technology (under my direction) in 2005 by an European board for postdoc formation in research labs.

EUROPEAN COMMISSION RESEARCH DIRECTORATE-GENERAL

Directorate D - The human factor, mobility and Marie Curie activities

Brussels, 4 August 2005- RTD-D5/AH D(2005) / 528708

Quality of the Hosts (Weight 15 / No Threshold) Mark (out of 5)

“The host institution is a top institution in the field of cognitive science. Its international connections are impressive. Mark 4,8”

IMPACT: Castelfranchi’s impact in these areas is visible in citation analysis:

H-index=56, (Google Scholar december 2016), i10-index 190, Citations 14006),

1 monograph with 780+ citations, 12 papers with 200+ citations. Importantly, these contributions were published and cited across several disciplinary areas, such as cognitive science (e.g. Behavioral and Brain Sciences), psychology (e.g. Cognitive Processing), computer science (e.g. Artificial Intelligence), philosophy (e.g. Synthese), and social sciences (e.g. International Review of Economics; Mind & Society; JASSS).

RELEVANT PUBLICATION:

- Most influential works (>>>)

Parisi, D. e Castelfranchi, C. Una definizione della psicologia cognitivista. In G. Kanizsa e P. Legrenzi (a cura di) *Psicologia della Gestalt e psicologia cognitivista*. Bologna: Il Mulino, 1978.

Castelfranchi, C., Miceli, M. e Parisi, D. Invidia. *Giornale Italiano di Psicologia*, VII, 1980, 95-119.

Parisi, D. e Castelfranchi, C. A goal analysis of some pragmatic aspects of language. In H. Parret, M. Sbisà and J. Verschueren (Eds.), *Possibilities and Limitations of Pragmatics*. Studies In Language Companion Series (Vol. 7). Amsterdam: John Benjamins, 1981.

Poggi, I., Castelfranchi, C. e Parisi D. Answers, Replies and Reactions. In H. Parret, M. Sbisà and J. Verschueren (Eds.), *Possibilities and Limitations of Pragmatics*. Studies In Language Companion Series (Vol. 7). Amsterdam: John Benjamins, 1981.

>> Vincent, J. and Castelfranchi, C. On the art of deception: How to lie while saying the truth. In H. Parret, M. Sbisà and J. Verschueren (Eds.), *Possibilities and Limitations of Pragmatics*. Studies In Language Companion Series (Vol. 7). Amsterdam: John Benjamins, 1981.

Castelfranchi, C. Scopi esterni. *Rassegna Italiana di Sociologia*, XXIII, 3, 1982.

Castelfranchi, C., Devescovi, A. e Burani, C. Understanding causal relations. In *Proceedings of the Conference on language, reasoning and inference: An interdisciplinary conference in cognitive science*. Edinburgh: University of

- Edinburgh, School of Epistemics, pp.2-6. 1982, April.
- Castelfranchi, C., Parisi, D. e Stock, O. Natural Language Project. *Sigart Newsletter*, 1982, January, 79-80. .
- Castelfranchi, C., Parisi, D. e Stock, O. Free order languages. An experimental lexicon based parser. *Ninth Conference of Computational Linguistics - COLING'82*, Praga, 1982.
- Castelfranchi, C., Parisi, D. e Stock, O. A lexicon based approach to sentence understanding. In R. Trappl (Ed.) *Proceedings of Cybernetics and System Research*, Amsterdam: North-Holland, 1982.
- Castelfranchi, C., Parisi, D. e Stock, O. Non-determinism and idiom recognition in a lexicon based syntactic parser for italian. *Proceedings of ECAI-82. European Conference on Artificial Intelligence*, Orsay, July, 1982.
- Stock, O., Castelfranchi, C. e Parisi, D. Wednesday: Parsing flexible word order languages. *Proceedings of the 1st meeting of the Association for Computational Linguistics - ACL. European Chapter*. Pisa, 1983.
- Castelfranchi, C., Parisi, D. *Mente e scambio sociale. Rassegna Italiana di Sociologia*, XXV, 1. 1984.
- Castelfranchi, C., Parisi, D. e Stock, O. Extending the expressive power of propositional nodes. In G. Guida and B. Bara (Eds.), *Natural language processing*. Amsterdam: North Holland, 1984.
- Castelfranchi, C. e Poggi, I. Communication: beyond the cognitive approach and the speech act theory. *Proceedings of International Conference on Pragmatics*, Amsterdam: Benjamins, 1985.
- Castelfranchi, C. e Stock, O., Concept, class, prototype: Unum an trinum. In R. Trappl (Eds.), *Cybernetics and Systems*, Reidel, 1986, 793-798.
- Parisi, D. e Castelfranchi, C. *La macchina e il linguaggio*. Torino: Boringhieri. 1987.
- Parisi, D., Castelfranchi, C. e Stock, O. Linguaggio naturale e processi cooperativi uomo-computer. In O. Stock (Eds.) *Interazione uomo-macchina intelligente*. Roma: Il Pensiero Scientifico, 1987.
- Castelfranchi, C. *Che figura. Emozioni e immagine sociale*. Il Mulino, Bologna, 1988.
- Castelfranchi, C., D'Aloisi, D. Stock, O e Tuozzi A. Propositions in a Hybrid Knowledge Representation System. In T. O'Shéa and Sgurev (eds.) *Artificial Intelligence III*, Elsevier, Amsterdam, 1988.
- Parisi, D. e Castelfranchi, C., Disambiguation in a lexically based sentence understanding system. In S. Small, G. Cottrell an M. Tanenhaus (eds.) *Lexical Ambiguity Resolution*, Morgan Kaufmann, San Mateo, Cal., 1988.
- Castelfranchi, C. Paradisi artificiali: prolegomeni alla simulazione della interazione sociale. *Sistemi Intelligenti*, I, 1, 123-165, 1989.
- Castelfranchi, C. e Cesta, A. No Control without Understanding: Problems in Knowledge Representation for Monitoring. *Proceedings of the Second European Meeting on Cognitive Science: Approaches to Process Control*, Siena, Italy, Oct 1989.
- Castelfranchi, C., Cesta, A., Conte, R. e Miceli, M. Conoscenze per l'interazione: aspetti sociali dell'agente autonomo intelligente. *Proceedings del 1° Congresso della Associazione Italiana per l'Intelligenza Artificiale (AI*IA'89)*, Trento, nov. 1989.
- Cesta, A., Castelfranchi, C., Conte, R. Using Flexible Pragmatic Structures for Augmenting Knowledge of Expert Systems, *Proceedings of CIL-89*, Barcellona.
- Cesta, A., Castelfranchi, C., Conte, R. e Miceli, M. Deep KR about Goal Achievement: Motivations for Pragmatic Structures. In A. Martelli e G. Valle *Computational Intelligence-88*, Elsevier, Amsterdam, 1989.
- Cesta, A., Romano, G. e Castelfranchi, C. Representation of Past Plans for Reusing. In I. Plander, *Artificial Intelligence and Information-Control Systems of Robots-89*, Elsevier, 1989.
- Conte, R. Castelfranchi, C. e Cesta, A. Natural Topic-Change in Plan-Recognition. *EPIA-89, Portuguese Conference of Artificial Intelligence*, Lisboa, Portugal, Sept. 1989.
- Conte, R., Castelfranchi, C., Cesta, A. e Miceli, M. Pragmatic Structures for Social Interaction, In M. C. Jackson, P. Keys, S. Cropper (eds.), *Operational Research and Social Sciences*, Plenum publ., London, 1989.
- Conte, R., Cesta, A. e Castelfranchi, C. How to Computationalize the Cognitive Agent? *Proceedings of Intelligent Management Systems*, Varna, Bulgaria, sept. 1989.

- >> Miceli, M. e Castelfranchi, C. A Cognitive Approach to Values, *Journal for the Theory of Social Behaviour*, vol. 19, 2, 169- 194, 1989.
- >> Castelfranchi, C. Social power: a point missed in Multi-Agent, DAI, and HCI. In Y. Demazeau and J. P. Muller, *Decentralized A. I.*, North-Holland, Amsterdam, 1990.
- Castelfranchi, C. e D'Aloisi, D. Hybrid Hierarchies: a love-hate relationship between ISA and SUPER. In R Trappl (ed.), *Cybernetics and Systems '90*, World Scientific, London, 1990.
- Castelfranchi, C. e D'Aloisi, D. A cognitive approach to knowledge revision *COGNITIVA-90*, AFCET, Paris, 1990.
- >> Castelfranchi, C. e Poggi, I. Blushing as a discourse: was Darwin wrong? In R. Crozier (ed.) *Shyness and Embarrassment: Perspective from Social Psychology*, Cambridge University Press, N. Y., 1990.
- Castelfranchi, C. e Poggi, I. Ingredienti per una macchina che si vergogna. In V. D'Urso (a cura di), *Imbarazzo, vergogna e altri affanni*. Cortina, Milano, 1990.
- Conte, R e Castelfranchi, C. Issues on the mental representation of norms. *Proceedings of the 2nd International Congress for Research on Activity Theory*. Lahti, Finland, 1990.
- Conte, R., Miceli, M., Castelfranchi, C. e Cesta, A. Planning knowledge for cognitive agents *COGNITIVA-90*, AFCET, Paris, 1990.
- Miceli, M e Castelfranchi, C. Evaluation: a bridge between knowledge and goals. *Proceedings of the Summer School on Programming 90*, Varna, Bulgaria, Maggio, 1990.
- Castelfranchi, C. Emozione e regolazione del comportamento. In T. Magri e F. Mancini (a cura di) *Emozione e conoscenza*, Editori Riuniti, Roma, 1991.
- Castelfranchi, C. Forse tu non pensavi che io Loica fossi. *Sistemi Intelligenti*, Il Mulino, (vol.)1 141-160, 1991.
- Castelfranchi, C. Altruismo, *Enciclopedia delle Scienze Sociali*, Istituto Treccani, vol. 1, pp. 137-145, Roma, 1991.
- Castelfranchi, C. e Conte, R. Le strutture della mente normativa. Problemi di rappresentazione mentale delle norme. In R. Conte (a cura di) *La norma. Mente e regolazione sociale*, Editori Riuniti, Roma, 1991.
- Castelfranchi, C. e Conte, R. Emergent functionalities among intelligent systems: Cooperation within and without minds. *AI and Society*, 1991.
- Conte, R. e Castelfranchi, C. Society out of minds and in the minds. In L. Gasser (ed.) *Knowledge and Action at Social & Organizational Levels*, AAAI Fall Symposium Series, 1991.
- Conte, R, Miceli, M e Castelfranchi, C. Limits and Levels of Cooperation. Disentangling various types of prosocial interaction. In Y. Demazeau & J. P. Muller (Eds.) *Decentralized AI - 2*, Amsterdam, Elsevier, 1991.
- Miceli, M. e Castelfranchi, C. A means for arguing. *Proceedings of AAAI Spring Symposium on Argumentation*, Palo Alto, 1991.
- Castelfranchi, C. No More Cooperation, Please! In Search of the Social Structure of Verbal Interaction. In Ortony, A, Slack, J. and Stock, O. (eds.) *A. I. and Cognitive Science Perspectives on Communication*. Heidelberg, Germany: Springer, 1992.
- Castelfranchi, C., Conte, R., Cesta, A. The Organization as a Structure of Negotiated and Non Negotiated Binds, in G. C. van der Veer, M. J. Tauber, S. Bagnara, M. Antalovits (Eds.), *Human Computer Interaction: Tasks and Organization. Proceedings of the Sixth European Conference on Cognitive Ergonomics (ECCE6)*, Balatonfured, Hungary, September 6-11, 1992.
- >> Castelfranchi, C., Miceli, M. e Cesta, A. Dependence relations among autonomous agents. In E. Werner, Y. Demazeau (Eds), *Decentralized A. I. - 3* , pp. 215-227, North Holland, Amsterdam, 1992. .
- Castelfranchi, C. Cooperazione e linguaggio: modelli di socialità? *Sistemi Intelligenti*, 4, 1, 1992, pp. 167-74.
- Castelfranchi, C. Autonomy. In *Atti del Workshop on Abstract Agent Architecture*, ENEA, Roma, febb. 1993.
- Castelfranchi, C. Solidarietà, mano invisibile e bene pubblico: elementi dell'ideologia corrente. In B. Cattarinussi (a cura di) *Altruismo e solidarieta'*. Milano Franco Angeli, 1993, 35-48.
- >> Castelfranchi, C. Social Commitment: from individual intentions to groups and organizations. In *ICMAS'95 First International Conference on Multi-Agent Systems*, AAAI-MIT Press, 1995, 41-49 (versione preliminare in *AAAI Workshop on "AI and Theory of Group and Organization"*, Washington, D. C., maggio 1993).
- >> Conte, R. & Castelfranchi, C. Norms as mental objects. From normative beliefs to normative goals. In *Proceedings of the 5th European Workshop on MAAMAW*, Neuchatel, Switzerland, August 24-28. 1993.

- D'Aloisi, D., Castelfranchi, C. Propositional and Terminological Knowledge Representations, *Journal of Experimental and Theoretical Artificial Intelligence*, 5 (3), 1993.
- Falcone, R., Castelfranchi, C. Plan Recognition for Explanation: a Chart-based system, *Proceedings of the Eighth International Conference of Applications of Artificial Intelligence in Engineering (AIENG-93)*, Toulouse, France, June, 1993.
- Falcone, R. e Castelfranchi, C. CHAPLIN: a chart based plan recognizer, *Proceedings of 13° Intern. Conf. of Avignone*, 1993.
- Castelfranchi, C. Survey. In *International Journal of Mental Health. Special issue on "Jobs and Psychosocial Rehabilitation: Italian Experiences and Models"*, 1994, vol.23,1, pp.86-90
- Castelfranchi C. Intelligenza Artificiale Distribuita. In O. Stock (a cura di) *L'Intelligenza Artificiale: aree, tendenze e prospettive*. F. Angeli, Milano, 1994.
- Castelfranchi C. Colpevol-mente: alcuni punti (interrogativi) sui sensi di colpa. In Castelfranchi C., R. D'Amico, I. Poggi, *Sensi di colpa: aspetti cognitivi, affettivi e relazionali*, Firenze, Giunti, 1994, pp. 59-100.
- Castelfranchi C. "Ma non dica idiozie!". Per un modello delle interazioni verbali al di la' della conversazione, in F. Orletti (a cura di) *Fra conversazione e discorso. L'analisi della interazione verbale*, Roma, La Nuova Italia Scientifica, 1994, pp. 143-170.
- Castelfranchi C. La psicologia come invitato di pietra al banchetto tra Logica ed Intelligenza Artificiale, in C. Cellucci, M.C. Di Maio, G. Roncaglia (a cura di) *Logica e Filosofia della scienza: problemi e prospettive*, Pisa, ETS, 1994, pp. 769-72
- Castelfranchi C., Conte R., Diani M. Paradossi cognitivi della democrazia e limiti all'azione del cittadino, in *Parolechiave*, 5, 1994, pp. 33-63.
- Castelfranchi, C., Falcone, R. (1994). Towards a theory of single-agent into multi-agent plan transformation. *The 3rd Pacific Rim International Conference on Artificial Intelligence*, Beijing, China, 16-18 agosto.
- Castelfranchi C. & Poggi I. Lying as Pretending to Give Information. In H. Parret (ed) *Pretending to Communicate*, New York, W.de Gruyter, 1994, pp. 276-91.
- Castelfranchi C. & Werner E. (eds.) *Artificial Social Systems*, Heidelberg, Springer-Verlag, LN/AI 830, 1994.
- Castelfranchi C. & Werner E., The MAAMAW spirit, Introduction to Castelfranchi C. & E. Werner (eds) *Artificial Social Systems*, Heidelberg, Springer-Verlag, LN/AI 830, 1994.
- Conte, R. & Castelfranchi, C. Mind is not enough. Precognitive bases of social interaction. In N. Gilbert (Ed.), *Proceedings of the Symposium on Simulating Societies 1992*, London, University College of London Press, 1994.
- Falcone, R., Castelfranchi, C. (1994). Plan Recognition: from Single-Agent to Multi-Agent plans. *Proceedings of MAAMAW '94*, 3-5 August, 1994.
- >> Castelfranchi, C. Guaranties for Autonomy in Cognitive Agent Architecture. In M. Wooldridge and N. Jennings (Eds.) *Intelligent Agents*. Springer. LNAI 890, 1995, 56-70.
- Castelfranchi, C. (1995). Self-awareness: Notes for a computational theory of intra-psyhic social interaction. In G. Trautteur (ed.), *Consciousness*, Napoli, Bibliopolis, pp. 55-80.
- Castelfranchi, C. e Conte, R. (1995) Distributed Artificial Intelligence and Social Science: Critical Issues. In G.M.P. O'Hare and N.R. Jennings (eds.), *Distributed Artificial Intelligence: An Introduction*, Wiley, Sixth Generation Computer Technology Series.
- Castelfranchi, C. e Falcone, R. (1995). "To say and to do" virtual actions in the structure and recognition of discourse plans with regard to practical plans, *INTERACT '95*, Lillehammer, Norway, 27-29 June, 1995, pp.103-108.
- Castelfranchi, C. e Falcone, R. (1995). From Single-Agent to Multi-Agent: Challenges for Plan Recognition Systems *IJCAI'95 workshop "The Next Generation of Plan Recognition Systems: Challenges for and Insight from Related Areas of AI"*, August 20.
- Castelfranchi C. & Muller, J-P. (1995). Introduction to *From Reaction to Cognition*. Heidelberg, Springer-Verlag, LN/AI 957, pp.3-13.
- Castelfranchi, C. & Muller, J-P. (eds.) (1995) *From Reaction to Cognition*. Heidelberg, Springer-Verlag, LNAI 957.

- >> Castelfranchi, C., Conte, R. e Paolucci, M. Normative reputation and the costs of compliance. In *Proceedings of the 3rd Simulating Societies Symposium*, Boca Raton, Florida, USA, 15-17 Settembre, 1995.
- Castelfranchi, C., D'Aloisi, D., e Giacomelli, F. (1995). A framework for dealing with belief-goal dynamics. In M. Gori and G. Soda (eds.), *Topics in Artificial Intelligence* (pp. 237-242). Berlin: Springer-Verlag.
- >> Conte, R. e Castelfranchi, C. (1995). *Cognitive and social action*. London, UCL Press.
- Conte, R. e Castelfranchi, C. Norms as mental objects. From normative beliefs to normative goals. In Castelfranchi C. e J.P. Muller (eds.) *From reaction to cognition*, (pp.186-199). Berlin, Springer Verlag, 1995.
- >> Conte, R. e Castelfranchi, C. (1995). A simulative understanding of norm functionalities in social groups. In R. Conte e N. Gilbert (eds.) *Artificial societies: The computer simulation of social life*, UCL Press, 1995, 252-267.
- Conte, R. e Castelfranchi, C. (1995). The treatment of norms in modelling rational agents: cognitive issues. In *Proceedings of ModelAge 95, General Meeting of Esprit Working Group 8319*, INRIA, Sophia Antipolis, France, 22-24, Gennaio, 1995.
- Dongha, P. e Castelfranchi, C. (1995). Rationality in Commitment. In M. M.Fehling, D. Perlis, M. Pollack, J. Pollock (eds.), *Proceedings of AAAI Fall Symposium on Rational Agency*, Boston, MA, nov. 1995.
- Castelfranchi, C.(1996). Reasons: Belief Support and Goal Dynamics. *Mathware & Soft Computing*, 3, pp. 233-47.
- Castelfranchi, C. (1996). Prescribed Mental Attitudes in Goal-Adoption and Norm-Adoption. In *Workshop "Norms, obligations, and Conventions" ICMAS'96*, Japan, August.
- Castelfranchi, C. (1996) Conflict Ontology. In *Proceedings ofECAI'97 Workshop on Conflicts*. Budapest.
- Castelfranchi, C., Falcone, R. (1996). "Levels of help, levels of delegation and agent modeling" *Agent Modeling Workshop: AAAI-96*. Portland, Oregon, August 4, 1996.
- Castelfranchi, C., Conte, R., Miceli, M. & Poggi, I. (1996). Emotions and Goals. In B. Kokinov (Ed.), *Perspectives on Cognitive Science*, vol. 2, NBU Series in Cognitive Science, pp. 130-145. Sofia: New Bulgarian University.
- Castelfranchi, C. & Miceli, M. (1996). Commentary on "Towards a Design-Based Analysis of Emotional Episodes", *Philosophy Psychiatry Psychology*, 3 (2), 129-133.
- Castelfranchi, C., Conte, R., Miceli, M. & Poggi, I. (1996). Emotions and Goals. In B. Kokinov (Ed.), *Perspectives on Cognitive Science*, vol. 2, NBU Series in Cognitive Science, pp. 130-145. Sofia: New Bulgarian University.
- Conte, R. e Castelfranchi, C. From conventions to prescriptions. Towards a unified theory of norms (Preliminary Report). IN J.L. Fiadeiro e P.Y. Schobbens (eds.) *Proceedings of the 2nd Workshop of the ModelAge Project*, Sesimbra, Portugal, January 15-17, 1996, pp. 67-81.
- Conte, R. e Castelfranchi, C. (1996) Simulating multi-agent interdependencies. A two-way approach to the micro-macro link. In U. Mueller & K. Troitzsch (eds) *Microsimulation and the social science*. Berlin, Springer Verlag, Lecture Notes in Economics.
- Dongha, P. e Castelfranchi, C. Modelling Rational Commitment, *Proceedings of ModelAge '96*. Sesimbra, Portugal, January 14-17, 1996.
- Falcone, R., Castelfranchi, C., (1996). Plan Recognition: from Single-Agent to Multi-Agent plans. in J.Perram & JP. Muller (eds.) *Distributed Software Agents and Applications*, Lecture Notes in Artificial Intelligence. Springer-Verlag pg.166-178.
- Veneziano, V. Conte, R. e Castelfranchi, C. Formation of partnerships: The mediatory role of dependence networks. *Mitteilungen aus den Arbeitskreisen, MASSIM-96*, Heft nr. 53, Ulm 5-6 März 1996.
- Castelfranchi, C. (1997). Principles of Individual Social Action. In Tuomela R. & Hintikka, G. (eds) *Contemporary Action Theory*. Kluwer.
- Castelfranchi, C. (1997) To Be Or Not To Be An Agent. In J. Muller, M. Wooldridge, N. Jennings (Eds.) *Intelligent Agents III*, Springer, Berlin, 37-41.
- Castelfranchi, C. (1997) Representation and integration of multiple knowledge sources: issues and questions. In Cantoni, Di Gesu', Setti e Tegolo (eds), *Human & Machine Perception: Information Fusion*. Plenum Press.
- Castelfranchi, C. Prevenzione. Tra pessimismo della ragione e ottimismo della volonta'. In R. Piccione e A. Grispieni (a cura di) *Prevenzione e salute mentale*. Carocci, Roma, 1997.

- Castelfranchi, C., Falcone, R. (1997). Delegation Conflicts, in M. Boman & W. Van de Velde (eds.) *Multi-Agent Rationality*, Lecture Notes in Artificial Intelligence, 1237. Springer-Verlag pp. 234-254.
- Castelfranchi C., Falcone R. (1997). From Task Delegation to Role Delegation, in M. Lenzerini (Editor), *AI*IA97: Advances in Artificial Intelligence*, Lecture Notes in Artificial Intelligence, 1321. Springer-Verlag pp.278-289.
- Castelfranchi, C., Conte, R. e Paolucci, M. Normative compliance and the costs of transgression. In Conte, R. e Chattoe, E. (eds.), *Evolving societies: the computer simulation of social systems*, 1997.
- Castelfranchi C., de Rosis F., Falcone R. (1997). Social Attitudes and Personalities in Agents, Socially Intelligent Agents, *AAAI Fall Symposium Series 1997*, MIT in Cambridge, Massachusetts, November 8-10, pp.16-21.
- Castelfranchi C., de Rosis F., Falcone R. and Pizzutilo, S. (1997). A Testbed for investigating personality-based multiagent cooperation, *Proceedings of the Symposium on Logical Approaches to Agent Modeling and Design*, Aix-en-Provence, August, pp.18-22.
- Falcone, R., Castelfranchi, C., "On behalf of ..": levels of help, levels of delegation and their conflicts, *4th ModelAge Workshop: "Formal Model of Agents"*, Certosa di Pontignano (Siena), 15-17 gennaio 1997.
- Miceli, M. e Castelfranchi C. (1997). Basic principles of psychic suffering: A preliminary account. *Theory & Psychology*, 7, 769-798.
- Miceli, M. e Castelfranchi, C. (1997). Sofferenza psichica: principi strutturali. *Sistemi Intelligenti*, 9, 7-45.
- Pistolesi, G. Veneziano, V., Castelfranchi, C. Global vs. Local Social Search in the Formation of Coalitions. In R. Conte, R. Hegselmann. P. Terna (eds.) *Simulating Social Phenomena*. Springer, Lectures Notes in Economics and Mathematical Systems 456, pp. 203-10, 1997.
- >> Castelfranchi, C. (1998). Modelling Social Action for AI Agents. *Artificial Intelligence*, 103, 1998, 157-182.
- Castelfranchi, C. Through the minds of the agents, *Journal of Artificial Societies and Social Simulation*, 1(1), 1998 <http://www.soc.surrey.ac.uk/JASSS/1/1/contents.html>
- Castelfranchi, C. Simulating with Cognitive Agents: The Importance of Cognitive Emergence. In J. Sichman, R. Conte, N. Gilbert (eds.) *Multi-Agent Systems and Agent-Based Simulation*. Springer, Berlin LNAI 1534, 1998, 26-44.
- Castelfranchi, C. Emergence and Cognition: Towards a Synthetic Paradigm in AI and Cognitive Science. In H. Coelho (Ed.) *Progress in Artificial Intelligence - IBERAMIA 98*, Springer, LNAI1484, Berlin, 1998 13-26.
- Castelfranchi, C. Towards an Agent Ontology: Autonomy, Delegation, Adaptivity. *AI*IA Notizie*. 11, 3, 1998; Numero speciale su "Agenti Autonomi Intelligenti", Associazione Italiana per l'Intelligenza Artificiale, Roma, 45-50.
- >> Castelfranchi, C. To believe and to feel: The case of "needs". In D. Canamero (ed.) *Proceedings of AAAI Fall Symposium "Emotional and Intelligent: The Tangled Knot of Cognition"* 1998, AAAI Press, 55-60.
- Castelfranchi, C. e Conte, R. Limits of economic rationality for agents and MA systems. *International Journal of Robotics and Autonomous Systems*, Special issue on *Multi-Agent Rationality*, Elsevier Editor, 24, 1998, 127-139.
- >> Castelfranchi C., Falcone R. (1998). Principles of trust for MAS: cognitive anatomy, social importance, and quantification, *Proceedings of the International Conference on Multi-Agent Systems (ICMAS'98)*, Paris, July, pp.72-79.
- Castelfranchi C., Falcone R. (1998). Social Trust: cognitive anatomy, social importance, quantification and dynamics. *Autonomous Agents '98 Workshop on "Deception, Fraud and Trust in Agent Societies"*, Minneapolis/St Paul, USA, May 9, pp.35-49.
- >> Castelfranchi, C., Falcone, R. (1998) Towards a Theory of Delegation for Agent-based Systems, *International Journal of Robotics and Autonomous Systems*, Special issue on *Multi-Agent Rationality*, Elsevier Editor, Vol. 24, Nos 3-4, , pp.141-157.
- Castelfranchi C., Falcone R., de Rosis F. (1998). Deceiving in GOLEM: how to strategically pilfer help, *Autonomous Agents '98; Workshop on "Deception, Fraud and Trust in Agent Societies"*, Minneapolis/St Paul, USA, May 9, pp.100-112.
- Binmore K., Castelfranchi C., Doran J., Wooldridge M. Rationality in Multi-Agent Systems. *The Knowledge Engineering Review*, Vol. 13, 3, 1998, 309-14.
- Conte, R. , Castelfranchi, C., Veneziano, V. The computer simulation of partnership formation. *Computational and Mathematical Organization Theory* 4:4, 1998, 293-315.
- de Rosis F., Castelfranchi C., Falcone R., Covino E. (1998) Bayesian Cognitive Diagnosis in Believable Multiagent Systems, *International Belief Revision Workshop*, Trento, Italy, May 30 - June 1, pp.24-38.

- Falcone, R., Castelfranchi, C. (1998). Formalizing Trust. *Atti del Sesto Convegno dell'Associazione Italiana di Intelligenza Artificiale AI*IA '98*, Padova, 23-25 settembre, pp. 241-245.
- Miceli, M. e Castelfranchi, C. (1998). Denial and its reasoning. *British Journal of Medical Psychology*. 71, 139-152
- Miceli, M. e Castelfranchi, C. (1998). How to Silence One's Conscience: Cognitive Defences Against the Feeling of Guilt. *Journal for the Theory of Social Behaviour*, 28, 3, 287-318.
- >> Sichman, J.S., Conte, R., Castelfranchi, C., Demazeau, Y. (1998). A Social Reasoning Mechanism Based On Dependence Networks. In M. Hunhs and M. Singh (Eds.) *Readings in Agents*. Morgan Kaufmann, S. Francisco, 416-21.
- Castelfranchi, C. (1999). Why Computers Will (Necessarily) Deceive Us and Each Other. *ETHICOMP'99 - The future of information society*, Roma, LUISS, Ottobre 1999 - Proceedings elettronici.
- Castelfranchi, C. (1999). Prescribed Mental Attitudes in Goal-Adoption and Norm-Adoption. *AI & Law*, Special Issue on Norms in MAS. 7, 1999, 37-50.
- Castelfranchi, C. La fallacia dello psicologo. Per una teoria degli atti finalistici non intenzionali. *Sistemi Intelligenti*, XI, 3, 1999 435-68.
- Castelfranchi, C. and F. de Rosis. Which User Model do we need, to relax the hypothesis of 'Sincere Assertion' in HCI? *UM99 Workshop on 'Attitudes, Personality and Emotions in User-Adapted Interaction'*. Banff, June 1999.
- Castelfranchi, C., Falcone, R. (1999). The Dynamics of Trust: from Beliefs to Action, *Autonomous Agents '99 Workshop on "Deception, Fraud and Trust in Agent Societies"*, Seattle, USA, May 1, pp.41-54.
- Castelfranchi, C., Falcone, R. (1999). "Basic Mental Attitudes of a Collaborating Agent: Cognitive Primitives for MAS". In F.J. Garijo and M. Boman (Eds.) *Multi-Agent System Engineering*. Springer, LNAI 1647, Berlin, pp.188-209.
- Castelfranchi, C., de Rosis, F. and Grasso, F. Deception and suspicion in medical interactions: towards a simulation of believable In: *Machine Conversations*. Y Wilks (Ed), Kluwer International Series in Engineering and Computer Science, Vol 511, 1999.
- >> Castelfranchi, C., Dignum, F., Jonker, c. Treur, J. Deliberative Normative Agents: Principles and Architecture. In *ATAL'99 - International Workshop on Agents Theories, Architectures, and Languages*, 1999.
- >> Conte, R. & Castelfranchi, C. From conventions to prescriptions. Towards a unified view of norms. *Artificial Intelligence & Law*, 1999, 3, 323-40.
- de Rosis, F. and Castelfranchi C. How can personality factors contribute to make agents more 'believable'? Proceedings of the Workshop on 'Behaviour Planning for Lifelike Characters and Avatars'. Sitges, 1999.
- Falcone, R., Castelfranchi, C. (1999). Levels of Delegation and Levels of Help for Agents with Adjustable Autonomy, 1999 *AAAI Spring Symposium on Agents With Adjustable Autonomy*, March 22-24, 1999, Stanford University. pp 25-32.
- Falcone, R., Castelfranchi, C. (1999). Founding Agents Adjustable Autonomy on Delegation Theory, *Atti del Sesto Congresso dell'Associazione Italiana di Intelligenza Artificiale (AI*IA-99)*, Bologna, 14-17 settembre, pp. 1-10.
- >> Conte, R., Castelfranchi, C., Dignum, F. Autonomous Norm Acceptance. In J. Mueller (ed) *Proceedings of the 5th International workshop on Agent Theories Architectures and Languages*, Paris, 4-7 July, 1999.
- Castelfranchi, C. Per una teoria pessimistica della mano invisibile e dell'ordine spontaneo. In Salvatore Rizzello (a cura di) *Organizzazione, informazione e conoscenza. Saggi su F.A. von Hayek*. Torino, UTET, 2000.
- Castelfranchi, C. (2000). Artificial Liars: Why Computers Will (Necessarily) Deceive Us and Each Other. *Ethics and Information Technology*, 2: 113-119, 2000, Kluwer, Dordrecht.
- Castelfranchi, C. (2000). Affective Appraisal vs Cognitive Evaluation in Social Emotions and Interactions. In A. Paiva (ed.) *Affective Interactions. Towards a New Generation of Computer Interfaces*. Heidelberg, Springer, LNAI 1814, 76-106.
- Castelfranchi, C. Through the Agents' Minds: Cognitive Mediators of Social Action. *Mind and Society*. Torino, Rosembergh, 2000 pp.109-140.
- >> Castelfranchi, C. (2000). Engineering Social Order. In A. Omicini, R. Tolksdorf, F. Zambonelli (Eds.) *Engineering Societies in the Agents World*. Springer, LNAI 1972, 2000, pp. 1-18.
- Castelfranchi, C. Conflicts Ontology. In R Dieng and H.J Mueller (Eds.), *Conflicts in Artificial Intelligence*. 2000, Kluwer, Dordrecht. 21-40.
- Castelfranchi C., Falcone R. (2000). Trust and Control: A Dialectic Link, *Applied Artificial Intelligence Journal*, Special

- Issue on "Trust in Agents" Part1, Castelfranchi C., Falcone R., Firozabadi B., Tan Y. (Editors), Taylor and Francis 14 (8).
- Castelfranchi C., Falcone R. (2000). Does control reduce or increase trust? A complex relationship, *Autonomous Agents 2000 Workshop on "Deception, Fraud and Trust in Agent Societies"*, Barcelona, Spain, June 4, pp.49-60.
- Castelfranchi C., Falcone R. Trust is much more than subjective probability: Mental components and sources of trust, *32nd Hawaii International Conference on System Sciences - Track on Software Agents*, Maui, Hawaii, 5-8 January 2000. Electronic Proceedings 12 pag.
- Castelfranchi, C. Falcone, R. Conflicts within and for collaboration. In C. Tessier, L. Chaudron, HJ. Muller (Eds.) *Conflicting Agents: Conflict Management in Multi-Agent Systems*, Kluwer, Dordrecht, 2000, 33-61.
- Castelfranchi C., Falcone R., Firozabadi, B., Tan, Y.H. (2000). Guest Editorial, *Applied Artificial Intelligence Journal*, Special Issue on "Trust in Agents" Part1, Castelfranchi C., Falcone R., Firozabadi B., Tan Y. (Editors), Taylor and Francis 14 (8).
- Castelfranchi C., Falcone R., Firozabadi, B., Tan, Y.H. (2000). Guest Editorial, *Applied Artificial Intelligence Journal*, Special Issue on "Trust in Agents" Part2, Castelfranchi C., Falcone R., Firozabadi B., Tan Y. (Editors), Taylor and Francis 14 (9).
- Falcone, R., Castelfranchi, C., (2000). Grounding Autonomy Adjustment on Delegation and Trust Theory, *Journal of Experimental and Theoretical Artificial Intelligence*, Vol.12, N°2, pp. 149-152.
- Falcone, R., Castelfranchi, C.(2000). Delegation-based conflicts between client and contractor, *Journal of Intelligent Automation and Soft Computing*, TSI Press, Vol.6, n°3 pp. 193-204.
- Falcone R., Castelfranchi C. (2000). Il paradigma ad agenti: l'IA torna alle sue sfide originarie (e al successo). *Lettera matematica*, Springer-Verlag Editore, n°36, pp.21-29.
- Falcone, R., Castelfranchi, C. (2000). Levels of Delegation and Levels of Adoption as the basis for Adjustable Autonomy, *Lecture Notes in Artificial Intelligence* n°1792, pp.285-296.
- Conte, R., Castelfranchi, C. and Pedone, R. The impossibility of Modelling Cooperation in PD-Game. In J.J.Meyer and P-Y. Schobbens (Ed.) *Formal Models of Agents*. Springer, Berlin, 2000, 74-89.
- de Rosis, F., Castelfranchi, C. and Carofiglio, V. On various sources of uncertainty in modeling suspicion and how to treat them. *Workshop on 'Deception, Fraud and Trust in Agent Societies'*, *Autonomous Agents 2000*. Barcelona, June 2000.
- de Rosis, F., Grasso, F., Castelfranchi, C. and Poggi, I. Modeling conflict-resolution dialogues. R Dieng and J Mueller (Eds.) *Conflicts in Artificial Intelligence*, Kluwer, Dordrecht. , 2000
- de Rosis, F., Covino, M., Falcone, R. and Castelfranchi C. Bayesian cognitive diagnosis in believable multiagent systems. In M A Williams and H Rott (Eds.) *Frontiers of Belief Revision*, in the 'Applied Logics' Series of Kluwer, Dordrecht, 2000.
- Conte, R., Castelfranchi, C. Game-Theory: A Science Without Goals. In P.Y. Schobbens (ed), Berlin, Springer, (2000)
- Miceli, M., Castelfranchi, C. (2000). The Role of Evaluation in Cognition and Social Interaction. In K. Dautenhahn (Ed.) *Human Cognition and Social Agent Technology*. John Benjamins, Amsterdam, 2000.
- Miceli, M. e Castelfranchi, C. (2000). Nature and mechanisms of loss of motivation. *Review of General Psychology*, 4, 238-263.
- Castelfranchi C. and Yao-Hua Tan (eds.). *Trust and Deception in Virtual Societies*, by Kluwer Academic Publishers, 2001.
- Castelfranchi C. Information Agents: The Social Nature of Information and the Role of Trust (invited contribution). In M. Klusch, F. Zambonelli (eds.), *Cooperative Information Agents*, V, Heidelberg, Springer, LNAI 2182, pp. 208-210.
- Castelfranchi, C. (2001). The theory of social functions. Challenges for multi-agent-based social simulation and multi-agent learning. *Journal of Cognitive Systems Research* 2, 5-38. Elsevier. <http://www.cogsci.rpi.edu/~rsun/si-mal/article1.pdf>
- Castelfranchi, C. (2001). Towards a Cognitive Memetics: Socio-Cognitive Mechanisms for Memes Selection and Spreading. *Journal of Memetics.- Evolutionary Models of Information Transmission*, 5. http://www.cpm.mmu.ac.uk/jom-emit/2001/vol5/castelfranchi_c.html
- Castelfranchi C., Falcone, R. (2001). Social Trust: A Cognitive Approach. In *Trust and Deception in Virtual Societies* by Castelfranchi C. and Yao-Hua Tan (eds), Kluwer Academic Publisher, pp. 55-90.
- Castelfranchi C. and Yao-Hua Tan (2001). The Role of trust and deception in virtual society. *Proceedings of 34th Hawaii International Conference on System Sciences* 2001.
- Castelfranchi, C., Falcone, R. and de Rosis, F. Deceiving in GOLEM: how to strategically pilfer help. In: C Castelfranchi, and Y. H. Tan (Eds) *'Trust and Deception in Virtual Societies'*. Kluwer Publ Co. 2001 pp. 91-110.

- Castelfranchi, C. and Lesperance, Y. (eds.) *Intelligent Agents VII: Theories, Architectures, Languages* - Springer, LNAI, Berlin 2001
- Conte, R. & Castelfranchi, C. Are Incentives Good Enough?. In R. Conte and C. Dellarocas (Eds.) *Social Order in Multi-Agent Systems*, *Lecture Notes in AI*, Berlin: Springer Verlag, 2001, 45-61.
- Falcone R., Castelfranchi C. (2001). Agenti e Sistemi Multi-Agente, in Ernesto Burattini e Roberto Cordeschi (eds.) *Intelligenza Artificiale. Manuale per le discipline della comunicazione*. Roma, Carocci Editore 2001, pp. 227-254.
- Falcone R., Castelfranchi C. (2001). How trust enhance and spread trust, *Autonomous Agents 2001. Workshop on "Deception, Fraud and Trust in Agent Societies"*, Montreal, Canada, May 29.
- Falcone R., Castelfranchi C. (2001). The Human in the Loop of a Delegated Agent: The Theory of Adjustable Social Autonomy. In *System, Man and Cybernetics (Part A: Computer and Humans)*, vol. 31, No. 5, Settembre 2001, pp. 406-418.
- Falcone R., Castelfranchi C., (2001). The Socio-cognitive Dynamis of Trust: Does Trust Create Trust? In R. Falcone, M. Singh, Y.H. Tan (eds.) *Trust in Cyber-societies. Integrating the Human and Artificial Perspectives*. Heidelbergh, Springer, LNAI 2246, pp. 55-72.
- Falcone R., Castelfranchi C. (2001). Tuning the collaboration level with autonomous agents: A principled theory. *Proceedings of IJCAI-01 Workshop on "Autonomy, Delegation, and Control: Interacting with Autonomous Agents"*, H. Hexmoor, C. Castelfranchi, R. Falcone, and M. Cox (Eds) Seattle August 6, 2001; pp. 52-57.
- Cecchi R., Conte R., Castelfranchi C. Fraud, Exchange and Cooperation. In *Cooperative Agents*, Schmidt ed., Berlin: Springer, 2001.
- Miceli M., Castelfranchi C. (2001). Further Distinctions Between Coping and Defense Mechanisms? In *Journal of Personality*, 69: 2, Aprile 2001, pp. 287-296.
- Miceli M., Castelfranchi C. (2001). Acceptance as a Positive Attitude. In *Philosophical Explorations*, Vol. IV, Nr. 2, Maggio 2001, pp. 111-134.
- >> Miceli, M. e Castelfranchi, C. (2002). The mind and the future: The (negative) power of expectations. *Theory & Psychology*. Vol.12 (3): 335-366
- Miceli, M. e Castelfranchi, C. (Eds.) (2002). Special issue on "Desires, Goals, Intentions, and Values: Computational Architectures", *Cognitive Science Quarterly*, 2, Issues 3-4.
- Miceli, M. e Castelfranchi, C. (2002). Modelling motivational representations. *Cognitive Science Quarterly*, 2, 233-247.
- >> Castelfranchi, C. Formalising the Informal? Dynamic social order, bottom-up social control, and spontaneous normative relations. *Journal of Applied Logic 1 (2003) 47-92*.
- >> Castelfranchi, C. The Micro-Macro Constitution of Power, *ProtoSociology*, An International Journal of Interdisciplinary Research Double Vol. 18-19, 2003 *Understanding the Social II – Philosophy of Sociality*, Edited by Raimo Tuomela, Gerhard Preyer, and Georg Peter.
- Castelfranchi, C. (2003). Grounding We-intentions in Individual Social Attitudes. In Matti Sintonen & Petri Ylikoski & Kaarlo Miller (eds), "*Realism in Action - Essays in the Philosophy of Social Sciences*". Kluwer Publisher.
- Castelfranchi, C., Lorini, E. (2003). Cognitive Anatomy and Functions of Expectations. In *Proceedings of IJCAI'03 Workshop on Cognitive Modeling of Agents and Multi-Agent Interactions*, Acapulco, Mexico, August 9-11, 2003.
- Castelfranchi, C. For A "Cognitive Program" Explicit Mental Representations For *Homo Oeconomicus*. In *Cognitive Processes and Economic Behaviour* edited by Nicola Dimitri, Marcello Basili & Itzhak Gilboa, Routledge, London, 2003.
- Castelfranchi, C. (2003). "Theory of social and cognitive action and communication", Invited Speaker at the 1st International Workshop on "Socio-Cognitive Grids: The Net as a Universal Human Resource" - TALEs of the Disappearing Computer (DC), 1st - 4th of June 2003, Santorini, Greece - invited talk.
- Castelfranchi C., Falcone R. Social trust Theory. In "*The Open Agent Society*" Jeremy Pitt (Ed.) John Wiley & Sons, Ltd (2003).
- Castelfranchi C., Falcone R. (2003). From Automaticity to Autonomy: The Frontier of Artificial Agents, in Hexmoor H, Castelfranchi, C., and Falcone R. (Eds), *Agent Autonomy*, Kluwer Publisher, pp.103-136.
- Castelfranchi, C., Giardini, F. (2003), "Silent Agents. Behavioral Implicit Communication for M-A Coordination and HMI". The Second Annual Symposium on Autonomous Intelligent Networks and Systems, June 30th - July 1st 2003, Menlo Park, CA, U.S.A. - On-line proceedings.

- Castelfranchi, C., Giardini, F. (2003). "Will Humans and Robots Ever Coordinate Each Other Via Behavioral Communication?", Symposium on Evolvability and Interaction: Evolutionary Substrates of Communication, Signalling, and Perception in the Dynamics of Social Complexity, Queen Mary University, October, 8th-10th 2003, London, England - University of Hertfordshire technical report, pp. 30-35.
- Castelfranchi, C., Tummolini, L. (2003). Positive and Negative Expectations and the Deontic Nature of Social Conventions, in *Proceedings of the 9th International Conference of Artificial Intelligence and Law (ICAIL 2003)*, ACM Press: 119-125.
- Castelfranchi C., Falcone R., Pezzullo G., (2003). Belief Sources for Trust: Some Learning Mechanisms, *Proceedings of the Sixth International Workshop on Trust, Privacy, Deception and Fraud in Agent Systems c/o AAMAS-03*, Melbourne Melbourne (Australia), July 14, pp.101-106.
- Castelfranchi C., Falcone R., Pezzullo G. (2003) .Cooperating through a Belief-based Trust Computation. In *Proceedings of the twelfth IEEE International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprise (WETICE-03)*, IEEE Computer Society, ISBN 0-7695-1963-6, pp. 263-268.
- Castelfranchi, C., Giardini, F., Lorini, E., Tummolini, L. (2003) The Prescriptive Destiny of Predictive Attitudes: from Expectations to Norms via Conventions, in R. Alterman, D. Kirsh (Eds) *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, Boston, MA.
- Giardini, F., Castelfranchi, C. (2003). Beyond Language: Implicit Communication. *Technical Report for European Project Eurocores OMLL: The Origins of Man, Language and Languages*.
- Falcone R., Castelfranchi C., Pezzullo G. (2003). Trust in Information Sources as Source of Trust: A Fuzzy Approach. In *Proceedings of AAMAS 2003*. Melbourne.
- Hexmoor, H., Castelfranchi, C. e Falcone R. (2003). A Prospectus on Agent Autonomy, in Hexmoor H., Castelfranchi, C., and Falcone R. (Eds), *Agent Autonomy*, Kluwer Publisher, pp.1-10.
- Miceli, M. & Castelfranchi, C. (2003). The plausibility of defensive projection: A cognitive analysis. *Journal for the Theory of Social Behaviour*, 33, 279-301.
- Miceli, M. & Castelfranchi, C. (2003). Crying: Discussing its basic reasons and uses. *New Ideas in Psychology*, 21, 247-273.
- Castelfranchi C. (2004). Trust mediation in knowledge management and sharing. In C. Jensen, S. Poslad, and T. Dimitrakos, editors, *Trust Management: iTrust 2004*, LNCS 2995, pp. 304–318. Springer, 2004.
- >> Castelfranchi C. (2004). Silent Agents. From Observation to Tacit Communication. Modelling Other Agents from Observations: MOO 2004 A one-day workshop to be held as part of the International Joint Conference on Autonomous Agents and Multi-Agent Systems July 19, 2004 URL: <http://www.cs.biu.ac.il/~galk/moo2004/>
- Castelfranchi C., Falcone R. (2004) Founding Autonomy: The Dialectics between (Social) Environment and Agent's Architecture and Powers, in the *Special Issue on Agents and Computational Autonomy, Lecture Notes on Artificial Intelligence 2969*, Springer-Verlag, Berlin, pp. 40-54.
- Falcone R., Castelfranchi C. (2004). Trust Dynamics: How Trust is influenced by direct experiences and by Trust itself; *Proceedings of the 3rd International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-04)*, New York, 19-23 July 2004, ACM-ISBN 1-58113-864-4, pages 740-747.
- Giardini F., Castelfranchi, C. (2004). Behavior Implicit Communication for Human-Robot Interaction, *Proceedings of the AAAI Fall Symposium 2004 (The intersection of Cognitive Science and Robotics: From Interfaces to Intelligence)*, pp. 91-96.- AAAI Fall Symposium Series 2005, pp. 32-37
- Tummolini, L., Castelfranchi C., Ricci, A., Viroli, M., Omicini, A., "Exhibitionists" and "Voyeurs" do it better: a Shared Environment for Flexible Coordination with Tacit Messages, in H. van Parunak, D. Weyns (Eds.) *AAMAS 04 Proceedings of the Workshop on Coordination in Emergent Societies (E4MAS 2004)*.
- Tummolini, L., Castelfranchi C., Ricci, A., Viroli, M., Omicini, A. What I See is What You Say: Coordination in a Shared Environment with Behavioral Implicit Communication, in G. Vouros (Ed.) *ECAI 04 Proceedings of the Workshop on Coordination in Emergent Societies (CEAS 2004)*.
- Falcone R., Pezzullo G., Castelfranchi, C., Calvi G. (2004). Why a cognitive trustier perform better: Simulating trust-based Contract Nets; *Proceedings of the 3rd International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-04)*, New York, 19-23 July 2004, ACM-ISBN 1-58113-864-4, pages 1392-1393.
- Falcone R., Pezzullo G., Castelfranchi, C., Calvi G. (2004). Trust-based Contract Nets; in *Proceedings of the 5th International Workshop on "Engineering Societies in the Agents' World" (ESAW-04)* October 20-22, Toulouse (F), pp. 145-150.
- Lorini, E., Castelfranchi, C. (2004). The role of epistemic actions in expectations. Accepted paper for Second Workshop of

Anticipatory Behavior in Adaptive Learning Systems 2004 (ABIALS 2004), Los Angeles, 17 July 2004.

Lorini, E., Castelfranchi, C. (2004). An expectations-based theory of defense. Accepted paper for First International Workshop on Safety and Security in Multi-agent systems, New York, 20 July 2004.

Omicini, A. Ricci, A. Viroli, M., Castelfranchi C., Tummolini, L. (2004) Coordination Artifacts: Environment Based Coordination for Intelligent Agents, in *Proceedings of the International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2004)*, New York, USA.

Paglieri, F., Castelfranchi, C. (2004). Argumentation and belief revision: on the two-sided nature of epistemic change. In: F. Grasso, G. Carenini, C. Reed (eds.), *Proceedings CMNA IV*, Valencia, pp. 5-12.

Tummolini, L., Castelfranchi C., The Cognitive and Behavioral Mediation of Institutions: towards an Account of Institutional Actions. Paper presented at the Fourth Conference on Collective Intentionality, October 13-15 2004, Siena.

Falcone R., Castelfranchi, C. (2005). An Overview of a Belief-Based Model of Trust. *Encyclopedia of Information Science and Technology*, Volume I-V, Idea Group Inc. January.

Paglieri, F., Castelfranchi, C. (2005). Arguments as belief structures: Towards a Toulmin,s layout of doxastic dynamics. *Argumentation XIX*, 2 (in press).

Paglieri, F., Castelfranchi, C. (2005). Arguing within and between minds: Belief dynamics in Toulmin's layout of argument. In: *D. Hitchcock (ed.), The uses of argument: Proceedings of OSSA 2005*, Hamilton, Canada.

Paglieri, F., Castelfranchi, C. (2005). Revising beliefs through arguments: bridging the gap between belief revision and argumentation in MAS. In: I. Rahwan, P. Moratis, C. Reed (eds.), *Argumentation in Multi-Agent Systems*, Springer-Verlag, Berlin.

Falcone R., Pezzulo G., Castelfranchi, C., Calvi G. (2005). Contract Nets for Evaluating Agent Trustworthiness, Special Issue on "Trusting Agents for trusting Electronic Societies" in *Lecture Notes in Artificial Intelligence*, 3577 (in press).

Lorini, E., Marzo, F., Castelfranchi (2005). A cognitive model of the altruistic mind. International Conference on Cognitive Economics, Sofia.

[Luca Tummolini](#), Cristiano Castelfranchi: Trace Signals: The Meanings of Stigmergy. [E4MAS 2006](#): 141-156

Cristiano Castelfranchi, [Rino Falcone](#), [Michele Piunti](#): Agents with Anticipatory Behaviors: To Be Cautious in a Risky Environment. [ECAI 2006](#): 693-694

Cristiano Castelfranchi: SILENT AGENTS: From Observation to Tacit Communication. [IBERAMIA-SBIA 2006](#): 98-107

[Emiliano Lorini](#), [Andreas Herzig](#), Cristiano Castelfranchi: Introducing *Attempt* in a Modal Logic of Intentional Action. [JELIA 2006](#): 280-292

[Giovanni Pezzulo](#), [Gianluca Baldassarre](#), [Martin V. Butz](#), Cristiano Castelfranchi, [Joachim Hoffmann](#): From Actions to Goals and Vice-Versa: Theoretical Analysis and Models of the Ideomotor Principle and TOTE. [SAB ABIALS 2006](#): 73-93

Cristiano Castelfranchi: Why We Need a Non-reductionist Approach to Trust. [iTrust 2006](#): 1-2

Cristiano Castelfranchi, [Rino Falcone](#), [Francesca Marzo](#): Being Trusted in a Social Network: Trust as Relational Capital. [iTrust 2006](#): 19-32

[Luca Tummolini](#), Cristiano Castelfranchi: The cognitive and behavioral mediation of institutions: Towards an account of institutional actions. [Cognitive Systems Research 7](#)(2-3): 307-323 (2006)

[Luca Tummolini](#), Cristiano Castelfranchi: Introduction to the Special Issue on Cognition, Joint Action and Collective Intentionality. [Cognitive Systems Research 7](#)(2-3): 97-100 (2006)

Cristiano Castelfranchi, [Paul Davidsson](#), [André Meyer](#), [Maria Fasli](#), [Jan Odelstad](#), [Guido Boella](#), [Leendert van der Torre](#), [Harko Verhagen](#): List of reviewers. [Computational & Mathematical Organization Theory 12](#)(2-3): 277 (2006)

[Emiliano Lorini](#), Cristiano Castelfranchi: The Unexpected Aspects of Surprise. [IJPRAI 20](#)(6): 817-834 (2006)

[Salima Hassas](#), [Giovanna Di Marzo Serugendo](#), [Anthony Karageorgos](#), Cristiano Castelfranchi: Self-

Organising Mechanisms from Social and Business/Economics Approaches. [Informatica \(Slovenia\)](#) 30(1): 63-71 (2006)

[Emiliano Lorini](#), [Nicolas Troquard](#), [Andreas Herzig](#), Cristiano Castelfranchi: Delegation and mental states. [AAMAS 2007](#): 153

[Paolo Turrini](#), [John-Jules Ch. Meyer](#), Cristiano Castelfranchi: Rational Agents That Blush. [ACII 2007](#):

314-325



[Giovanni Pezzulo](#), [Gianguglielmo Calvi](#), Cristiano Castelfranchi: DiPRA: Distributed Practical Reasoning Architecture. [IJCAI 2007](#): 1458-1463

[Michele Piunti](#), Cristiano Castelfranchi, [Rino Falcone](#): Surprise as Shortcut for Anticipation: Clustering Mental States in Reasoning. [IJCAI 2007](#): 507-512

[Michele Piunti](#), Cristiano Castelfranchi, [Rino Falcone](#): Expectations driven approach for Situated, Goal-directed Agents. [WOA 2007](#): 104-111

[Giovanni Pezzulo](#), [Martin V. Butz](#), Cristiano Castelfranchi, [Rino Falcone](#): The Challenge of Anticipation, A Unifying Framework for the Analysis and Design of Artificial Cognitive Systems [Springer 2008](#)

[Rino Falcone](#), Cristiano Castelfranchi: Generalizing Trust: Inferencing Trustworthiness from Categories. [AAMAS-TRUST 2008](#): 65-80

[Emiliano Lorini](#), [Rino Falcone](#), Cristiano Castelfranchi: Trust within the Context of Organizations: A Formal Approach. [Formal Aspects in Security and Trust 2008](#): 114-128

[Andreas Herzig](#), [Emiliano Lorini](#), [Jomi Fred Hübner](#), [Jonathan Ben-Naim](#), Cristiano Castelfranchi, [Robert Demolombe](#), [Dominique Longin](#), [Laurent Vercouter](#): Prolegomena for a Logic of Trust and Reputation. [NORMAS 2008](#): 143-157

[Emiliano Lorini](#), [Michele Piunti](#), Cristiano Castelfranchi, [Rino Falcone](#), [Maria Miceli](#): Anticipation and Emotions for Goal Directed Agents. [The Challenge of Anticipation 2008](#): 135-160

[Giovanni Pezzulo](#), [Martin V. Butz](#), Cristiano Castelfranchi: The Anticipatory Approach: Definitions and Taxonomies. [The Challenge of Anticipation 2008](#): 23-43

[Giovanni Pezzulo](#), [Martin V. Butz](#), Cristiano Castelfranchi, [Rino Falcone](#), [Gianluca Baldassarre](#), [Christian Balkenius](#), [Alexander Förster](#), [Maurice Grinberg](#), [Oliver Herbort](#), [Kiril Kiryazov](#), [Boicho N. Kokinov](#), [Birger Johansson](#), [Emilian Lalev](#), [Emiliano Lorini](#), [Carlos Martinho](#), [Maria Miceli](#), [Dimitri Ognibene](#), [Ana Paiva](#), [Georgi Petkov](#), [Michele Piunti](#), [Vin Thorsteinsdottir](#): Endowing Artificial Systems with Anticipatory Capabilities: Success Cases. [The Challenge of Anticipation 2008](#): 237-254

[Giovanni Pezzulo](#), [Martin V. Butz](#), Cristiano Castelfranchi, [Rino Falcone](#): Introduction: Anticipation in Natural and Artificial Cognition. [The Challenge of Anticipation 2008](#): 3-22

[Carles Sierra](#), Cristiano Castelfranchi, [Keith S. Decker](#), [Jaime Simão Sichman](#): 8th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009), Budapest, Hungary, May 10-15, 2009, Volume 1 [IFAAMAS 2009](#)

[Carles Sierra](#), Cristiano Castelfranchi, [Keith S. Decker](#), [Jaime Simão Sichman](#): 8th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009), Budapest, Hungary, May 10-15, 2009, Volume 2 [IFAAMAS 2009](#)

Cristiano Castelfranchi: Tacitly Communicating with Our Intelligent Environment via Our Practical Behavior and Its Traces. [Web Intelligence/IAT Workshops 2009](#): 323-326

Castelfranchi, C. Intra & inter-network dynamics, and the 'semantics' of the links. In BWSS 2010, Capes, Brazil,

Castelfranchi, C. Bye Bye Agents? Not. IEEE INTERNET COMPUTING, "Peering", 2010, pp. 93-96; <http://www-cdr.stanford.edu/~petrie/online/peer2peer/index.html>

Castelfranchi, C. Grounding Social Action in Mental Representations. "Advances in Cognitive Science: Learning, Evolution, and Social Action". Proceedings of IWCogSc-10- ILCLI, Universidad del Pais Vasco, 2010, pp. 93-112

Paglieri, Fabio and Castelfranchi, Cristiano (2010) 'Why argue? Towards a cost-benefit analysis of argumentation', *Argument & Computation*, 1: 1, 71 – 91

Castelfranchi, C. and Falcone, R. *Trust Theory. A Socio-Cognitive and Computational Model*. Wiley, UK, 2010.

Castelfranchi, C. Grounding Organizations in the Minds of the Agents Chapter X in Virginia Dignum (ed) Handbook of Research on Multi-Agent Systems: Semantics and Dynamics of Organizational Models, **Information science reference** - Hershey • New York

Castelfranchi C., Pezzulo G., Tummolini L. [*Behavioral Implicit Communication \(BIC\): Communicating with Smart Environments via our Practical Behavior and Its Traces*](#). In: International Journal of Ambient Computing and Intelligence, vol. 2 (1) pp. 1 - 12. IGI Global, 2010.

Castelfranchi, C. The "Logic" of Power. Hints on How my Power Becomes his Power. *SNAMAS-AISB'11*, 2011. pages 3-9.

Castelfranchi, C. (2011) È possibile una mente altruista? In "Altruismo e comportamento prosociale. Temi e prospettive a confronto" A cura di: [Stefano Boca](#), [Costanza Scaffidi Abbate](#) Franco Angeli Editore

Falcone, Rino, Piunti, Michele, Venanzi, Matteo and Castelfranchi, Cristiano (2011) From Manifesta to Krypta: The Relevance of Categories for Trusting Others. *ACM Transactions on Intelligent Systems and Technology.*, 4, (2), 1-24.

MARIA MICELI AND CRISTIANO CASTELFRANCHI Forgiveness: A Cognitive-Motivational Anatomy *Journal for the Theory of Social Behaviour* •••• 0021-8308, 2011

Tummolini L., Castelfranchi C. [L'incorporazione della proprietà : il problema delle istituzioni tra scienze cognitive sociali e scienze sociali cognitive](#). In: Sistemi Intelligenti, vol. 2 pp. 345 - 357. La cognizione sociale. A. Borghi, F. Caruana (eds.). il Mulino, 2011.

C Castelfranchi Goals, the true center of cognition, in **The goals of cognition** (edited by F. Paglieri, L. Tummolini, R. Falcone, M. Miceli) College Publications (London), **2012**

Castelfranchi, C. (2012b) Ascribing minds. *Cognitive Processing*. 13, 415-425

Castelfranchi C., Piunti M., Ricci A., Tummolini L. [Aml Systems as Agent-Based Mirror Worlds: Bridging Humans and Agents through Stigmergy](#). In: Agents and Ambient Intelligence - Achievements and Challenges in the Intersection of Agent Technology and Ambient Intelligence. pp. 17 - 31. T. Bosse (ed.). (*Ambient Intelligence and Smart Environments*, vol. 12). Amsterdam: IOS Press, 2012.

Andreas Herzig¹, Emiliano Lorini¹, Jomi F. Hu'bner², Jonathan Ben-Naim¹, Olivier Boissier², Cristiano Castelfranchi³, Robert Demolombe¹, Dominique Longin¹, Laurent Perrussel¹, and Laurent

Vercouter2 **Prolegomena for a logic of trust and reputation**
(NORMAS)

Miceli, M. & Castelfranchi, C. Consciousness in interaction. The role of the natural and social context in shaping consciousness. pp. 151 - 177. F. Paglieri (ed.). (Advances in Consciousness Research, vol. 86). Amsterdam: John Benjamins, 2012.



C Castelfranchi "Goal-directness" In: *Encyclopedia of Philosophy and the Social Sciences*. pp. 900 - 902. B. Kaldis (ed.). London: SAGE Publications Inc., 2013.

Cristiano Castelfranchi (2013) Minds as social institutions *Phenom Cogn Sci* 2013 DOI 10.1007/s11097-013-9324-0

C Castelfranchi *Cognitivizing Norms. Norm Internalization and Processing*. In *Informatica e Diritto*, Vol. XXII (2013), n. 1; Special Issue in "Law and Computational Social Science", S. Faro & N. Lettieri (Eds.); pp. 75-98.

Castelfranchi *Collective Unintelligence: The Ground of Society*
Extended Abstract for talk at ENSO Meeting: European Network on Social ontology, Helsinki, 2013

Castelfranchi C., Tummolini L. *I dedali interni della cognizione astratta*. In: *Giornale Italiano di Psicologia*, vol. XXXV (1) pp. 63 - 66. Il Mulino, 2013.

Andrighetto G., Castelfranchi C. Norm Compliance. The Prescriptive Power of Normative Actions. In: *Paradigmi*, vol. 2 pp. 120 - 135. Angeli, 2013

Andrighetto G., Castelfranchi C., Mayor E., McBreen J., Lopez-Sanchez M., Parsons S. *(Social) Norm Dynamics*. In: *Normative Multi-Agent Systems*. pp. 1 - 26. G. Andrighetto, G. Governatori, P. Noriega, L. van der Torre (eds.). (Dagstuhl Follow-Ups, vol. 4). Wadern, Germany: Dagstuhl Publishing, 2013.

Andrighetto G., Castelfranchi C. *Introduction to chapter III: Norms*. pp. 169 - 170. S. Ossowski (ed.). Springer, 2013

Francesca Marzo, Cristiano Castelfranchi (2013) Trust as Individual Asset in a Network: A Cognitive Analysis. In *Organizational Change and Information Systems. Working and Living Together in New Ways* Ed. Paolo Spagnoletti. Lecture Notes in Information Systems and Organisation Volume 2 2013 Pages 167-175

Tummolini L., Andrighetto G., Castelfranchi C., Conte R. *A convention or (tacit) agreement betwixt us: on reliance and its normative consequences*. In: *Synthese*, vol. 190 (4) pp. 585 - 618. Springer, 2013.

Castelfranchi, C. (2013) I gradi della mente: Probabilità * Piacere .Atti del AISC 2013- Convegno Associazione Italiana Scienza Cognitiva

Castelfranchi C. (2013) Mind in Degrees. The Quantitative Dimension of Mental Attitudes. Proceedings AI*IA, Springer

Roma 30 gennaio 2017

Prof. Cristiano Castelfranchi

“Ai sensi e per gli effetti del DLGS 196/2003 autorizzo il destinatario ad utilizzare il presente curriculum ed i dati personali in esso contenuti per le finalità e secondo le informazioni ricevute

Autocertificazione dei dati ai sensi dell’art.76 DPR 445/2000”