Self-organisation of Knowledge in Molecules of Knowledge

Stefano Mariani

DISMI
Università degli Studi di Modena e Reggio Emilia

Seminar
Autonomous Systems 16/17
June 7th, 2017

Outline

1. Introduction
2. $\mathcal{M}$olecules of $\mathcal{K}$nowledge
3. $\mathcal{M}o\mathcal{K}$ Pillars
4. Conclusion & Outlook

Outline

Context, Motivation, and Goal I

- ✓ Modern ICT systems go beyond Turing Machine like computation [Tur39] ⇒ *computation* = *algorithm* + *interaction* [Weg97]
- ⇒ How to manage interactions? ⇒ coordination models [MC94]
- ! *Open*, highly *dynamic*, and (mostly) *unpredictable* systems present novel challenges demanding innovative coordination approaches

> We deal with coordination issues in such a sort of systems by leveraging chemical-inspired and situated approaches, to promote *self-organisation*

Context, Motivation, and Goal II

✓ Socio-Technical Systems (STS) and Knowledge-Intensive Environments (KIE) combine processes, technologies, and *people*'s skills [Whi06] to handle large repositories of *information* [Bha01]

⇒ Managing their interaction space is of paramount importance for both functional and *non-functional* properties

! Engineering coordination mechanisms and strategies is far from trivial ⇒ *unpredictability* of agents' behaviour, *pace* of interactions, . . .

We integrate Behavioural Implicit Communication (BIC) in our approach, taming unpredictability to promote *anticipatory coordination*

Context, Motivation, and Goal III

✓ Data-driven approaches to coordination [DPHW05], e.g. tuple space based [Gel85] ⇒ coordinate interacting agents by managing *access to information*

⇒ Why to view data as *passive*, "dead" things to run algorithms upon in the traditional I/O paradigm?

We propose $\mathcal{M}$olecules $\mathcal{o}$f $\mathcal{K}$nowledge ($\mathcal{M}o\mathcal{K}$) as an innovative coordination model for *self-organising knowledge management*, interpreting *information as a living entity*

Outline

Outline

Overview I

Molecules of Knowledge (MoK) is a coordination model for *self-organisation of knowledge* in knowledge-intensive STS [MO13b]

- MoK promotes the idea that *data is alive* [CTZ02, ZOA+15], spontaneously interacting with other information and its prosumers (producer + consumer)

⇒ MoK pursues two main goals

- ✓ self-aggregation of information into meaningful heaps, possibly reifying relevant knowledge previously hidden
- ✓ spontaneous diffusion of information toward (potentially) interested agents

Overview II

- A $\mathcal{MoK}$-coordinated system is a network of information containers (*compartments*), in which sources of information (*seeds*) continuously and spontaneously inject atomic information pieces (*atoms*)...
- ...which may aggregate into composite information chunks (*molecules*), diffuse to neighbouring compartments, lose relevance as time flows, gain relevance when exploited, and the like ...
- ...according to decentralised and spontaneous processes dictating how the system evolves (*reactions*), influenced by agents' actions (*enzymes*) and their side effects (*traces*) ...
- ...which are transparently, and possibly unintentionally, caused by human or software agents (*catalysts*) while performing their activities

Core Abstractions I

Atoms *atomic units of information*, representing data along with its meta-data, decoorated with a concentration value resembling relevance

$$\texttt{atom}(\mathit{Src}, \mathit{Content}, \mathit{Meta\text{-}info})_c$$

Seeds *sources of information*, representing data sources as the collection of information they may make available

$$\texttt{seed}(\mathit{Src}, \mathit{Atoms})_c$$

Molecules *composite units of information*, representing collections of (semantically) related information

$$\texttt{molecule}(\mathit{Atoms})_c$$

Core Abstractions II

Catalysts *knowledge workers*, representing agents undertaking (epistemic) actions

$$Catalyst = (\alpha \dagger [\![\cdot]\!]_i).Catalyst$$

$$\alpha \in \{\texttt{share(}\mathit{Reactant}\texttt{)} \mid \texttt{mark(}\mathit{Reactant}\texttt{)} \mid \texttt{annotate(}\mathit{Reactant}\texttt{)} \mid \texttt{connect(}\mathit{Reactant}\texttt{)} \mid \texttt{harvest(}\mathit{Reactant}\texttt{)}\}$$

Enzymes *reification of actions*, representing the epistemic nature of actions and their context, enabling catalysts' to influence knowledge evolution

$$\texttt{enzyme(}\mathit{Species}\texttt{,}\mathit{s}\texttt{,}\mathit{Reactant}\texttt{,}\mathit{Context}\texttt{)}_c$$

Traces *reification of actions' (side) effects*, representing any (side) effect due to the action but not as its intentional primary effect

$$\texttt{trace(}\mathit{Msg}\texttt{,}\mathit{Context}\texttt{,}\mathit{Subject}\texttt{)}_c$$

Core Abstractions III

Perturbations *reactions to actions' side effects*, representing the computational functions enacted in response to agents' (inter-)actions and their side effects

`perturbation(`P`,`$Subject$`)`

P ::= `attract` | `repulse` | `approach` | `drift-apart` | `strengthen` | `weaken` | `boost` | `wane`

Reactions *knowledge dynamics processes*, representing the spontaneous computational processes supporting (meta-)information handling and evolution, as well as knowledge inference, discovery, and sharing, driven by (semantic) similarity of information

Core Abstractions IV

Compartments *knowledge containers*, representing the computational abstraction responsible for handling information lifecycle, provisioning data to agents, and executing reactions

$$Compartment = [\![\mathtt{Seeds}, \mathtt{Atoms}, \mathtt{Molecules}, \mathtt{Enzymes}, \mathtt{Traces}, \mathtt{Reactions}]\!]$$

Membranes *interaction channels*, representing the communication abstraction enabling 1 : 1 exchange of information, while defining the notions of locality and neighbourhood

$$Compartment_i \asymp Compartment_j$$

Reactions in a Nutshell I

$\mathcal{M}o\mathcal{K}$ reactions

- ✓ Reactions are chemical-like coordination laws executed according to *dynamic rate expressions* [Mar13]
  - ⇒ awareness of contextual information which may affect reactions application
  - ⇒ adaptiveness to external influences put by interacting agents
- ✓ The rationale driving reactions application is *(semantic)* similarity between reactant *templates* and *actual* reactants
  - according to $\mathcal{F}_{\mathcal{M}o\mathcal{K}}$ similarity measure

Reactions in a Nutshell II

Injection generates atoms from seeds

Aggregation ties together *(semantically)* related atoms, or molecules, into molecules

Diffusion moves atoms, molecules, and traces among neighbouring compartments

Decay decreases relevance of atoms, molecules, enzymes, and traces

Reinforcement increases relevance of atoms and molecules according to catalysts' *(inter-)actions*

Deposit generates traces from enzymes

Perturbation carries out the processes reacting to (side) effects of (interaction) activities undertaken by catalysts

(Inter-)actions in a Nutshell I

From Actions to Perturbations [MO15]

1. Catalysts' actions transparently release enzymes
   - ! each action ⇒ one *Species* of enzyme
2. Enzymes spontaneously and temporarily deposit traces
   - ! each enzyme ⇒ different traces ⇒ different perturbation actions
3. Traces diffuse to neighbouring compartments to apply perturbation actions
   - ! depending on availability of matching reactants and contextual information
4. Perturbation actions have different effects based on the trace they originate from and the current system state
   - ! different *Msg* + *Context* ⇒ different behaviour

(Inter-)actions in a Nutshell II

Catalysts' actions

`share` any action adding information to the system — posting information, sharing someone else's, . . .

`mark` any action marking information as relevant or not — liking a post, voting a question/answer, bookmarking a publication, . . .

`annotate` any action attaching information to other information — commenting posts, replying to comments, answering questions, and . . .

`connect` any action adding relationships with sources of information — adding friends, following people or posts, . . .

`harvest` any action acquiring knowledge — all kinds of search actions

(Inter-)actions in a Nutshell III

MoK traces

MoK interprets (inter-)actions according to Behavioural Implicit Communication (BIC) theory [CPT10]

- ! communication occurs (unintentionally) through *practical behaviour*
- ⇒ actions themselves, along with traces, *become the message* [MO13a]
- ✓ tacit messages, reified by MoK traces, describe these kind of messages

MoK exploits tacit messages through perturbation actions [MO15]

- ! leveraging mind-reading and signification abilities ascribed to agents and to the *computational environment*
- ✓ enabling anticipatory coordination according to the *ever-changing* needs of users

(Inter-)actions in a Nutshell IV

MoK perturbations

`approach/repulse` *facilitating/impeding interactions* between compartments whose agents interact more often than others

`attract/drift-apart` bringing to / taking from the compartment where the action took place *information (dis)similar* to the one target of the original action

`boost/wane` *increasing/decreasing rate* of specific MoK reactions to improve MoK coordinative behaviour

`strengthen/weaken` *increasing/decreasing relevance* of information (un)related to the one target of the original action

Matchmaking in a Nutshell

! $\mathcal{M}o\mathcal{K}$ needs a similarity measure for matchmaking

⇒ so as to promote *content-based* aggregation, reinforcement, diffusion, and perturbation

✓ $\mathcal{F}_{\mathcal{M}o\mathcal{K}}$ function represents the *fuzzy matchmaking* mechanism measuring similarity between atoms, molecules, etc.

⇒ *text-mining related* measures are exploited, e.g., cosine similarity, euclidean distance, average quadratic difference, . . .

! $\mathcal{F}_{\mathcal{M}o\mathcal{K}}$ depends on information representation

- e.g., for documents and excerpts of documents, experimented techniques include vector-spaces, key-phrases extraction, concept-based, . . .

MoK in Action: Interaction-driven Clustering

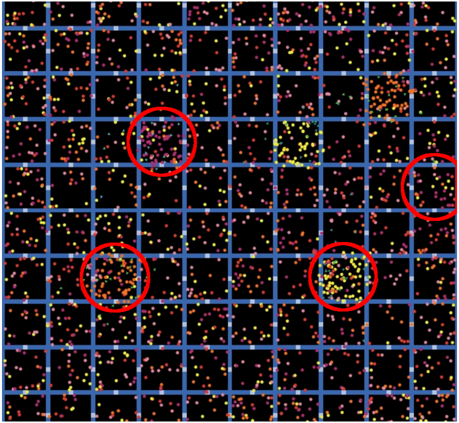


Figure: Whereas atoms and molecules are initially randomly scattered across compartments, as soon as catalysts interact clusters appear by emergence, thanks to *BIC-driven self-organisation*. Whenever new actions are performed by catalysts, MoK adaptively re-organises the spatial configuration of information so as to better tackle the new coordination needs

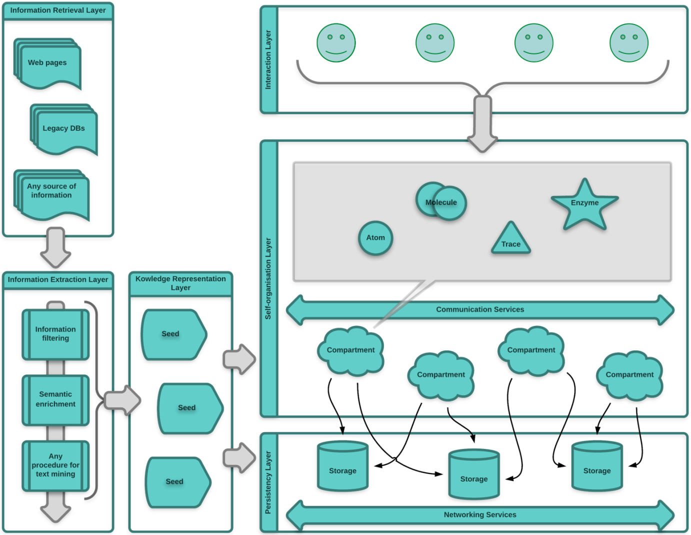
Outline

Premises

A comprehensive $\mathcal{M}o\mathcal{K}$ ecosystem would feature:

- ✗ automatic *information retrieval* and extraction
- ✓ automatic *semantic enrichment* of unstructured text
- ✗ graph and document oriented storage layer
- ✓ networking and communication facilities such as automatic *discovery* of compartments, *dynamic topology* re-configuration, gossiping, adaptive *routing*
- ✓ automatic *knowledge inference* and *discovery*, based on semantics
- ✗ interaction layer supporting behavioural implicit communication mechanisms to *assist* and *drive* automatic knowledge inference and discovery

MoK Ecosystem Architecture



Development Overview I

Networking

Asynchronous, channel-based services supporting automatic *discovery* of compartments, *dynamic re-configuration* of the network topology upon (dis)connections, point-to-point and multicast communication based on message passing

Communication

Gossiping algorithm based on probabilistic recursive multicast, *adaptive routing* for targeted communications between compartments, tolerant to dynamic network re-configurations

Development Overview II

Information harvesting

Natural Language Processing module for *part-of-speech tagging* and information extraction—automatic construction of atoms and molecules

Knowledge discovery

Similarity-based and *semantic-driven* aggregation of information based on WordNet synsets ontology

Still many open issues and improvement points:
feel free to join *MoK* dev team :)

Outline

Outline

Chemical Reactions as Coordination Laws

- LINDA model [Gel85] ⇒ simple yet expressive model for *fully uncoupled* coordination in *distributed* systems [Cia96]
- Socio-technical systems ⇒ uncertainty, unpredictability, adaptiveness
  - ✓ *unpredictability*, *uncertainty* ⇒ stochastic decision making
  - ✓ *adaptiveness* ⇒ programmability of the coordination machinery
  - ⇒ Biochemical tuple spaces [VC09], SAPERE [ZCF+11], ...
- Survey regarding bio-inspired design patterns [FMMSM+12]
  - ✓ [Nag04, DWH07, FMSM12, FMDMSA11, TRDMS11, VCMZ11]

1. Mechanism ⇒ artificial chemical reaction ⟺ coordination law
2. Evolution of the resulting "chemical solution" (coordination process) is simulated [Mar14]
   - ✓ different custom kinetic rates ⇒ different *emergent behaviours*

Probabilistic Coordination Primitives

- Uniform coordination primitives[1] (`uin`, `urd`) are *specialisations* of LINDA getter primitives featuring *probabilistic non-determinism* in returning matching tuples
- Uniform primitives feature global properties

  space: LINDA returns tuples *independently* of others, uniform primitives return tuples based on *relative multiplicity*

  time: sequences of LINDA operations exhibit no properties, sequences of uniform operations exhibit *uniform distribution*

- *Bio-inspired mechanisms* implemented on top of uniform primitives ⇒ behavioural expressiveness of uniform primitives [MO14]

[1] First mentioned in [GVCO07]

Outline

From A&A to Computational Smart Environments I

There is a gap in current approaches to STS engineering [SS00], which can be closed by dealing with

mutual awareness
: as the basis for *opportunistic*, ad hoc alignment and improvisation, which ensure *flexibility*

coordinative artefacts
: *encapsulating* those portions of the coordination responsibilities that is better to *automatise*

From A&A to Computational Smart Environments II

- Activity Theory (AT) is a social psychological theory for conceptualising human activities
  - ⇒ the A&A meta-model [ORV08] as a *reference framework* for designing the computational part of a STS for knowledge management
- Cognitive stigmergy [ROV+07] is a first generalisation of *stigmergy* where traces are amenable of a *symbolic interpretation*
  - ⇒ cognitive stigmergy directly supports both awareness and peripheral awareness in socio-technical systems
- Behavioural Implicit Communication (BIC) is a cognitive theory of communication [Cas06], where tacit messages describe the kind of messages a practical action (and its traces) may *implicitly* send to its observers [CPT10]
  - ⇒ BIC provides a sound cognitive and social model of action and interaction for both human agents and computational agents

From A&A to Computational Smart Environments III

BIC seem to provide mutual awareness, while *coordination artefacts* the required coordinative capabilities, paving the way toward computational smart environments [TCR+05]

Outline

Conclusion & Outlook

- We are in the perfect spot to start a paradigm shift toward self-organising knowledge, where:
  - user-centric adaptiveness of knowledge discovery processes is the foremost goal
  - measures and algorithms exploited for knowledge discovery, inference, management, and analysis natively account for users' goals
  - seamlessly scale up/down/out/in naturally, being operating on the assumption that only local-information is available consistently

Thanks for your attention

References I

Ganesh D. Bhatt.
Knowledge management in organizations: Examining the interaction between technologies, techniques, and people.
Journal of Knowledge Management, 5(1):68–75, 2001.

C Castlefranchi.
From conversation to interaction via behavioral communication: For a semiotic design of objects, environments, and behaviors.
Theories and practice in interaction design, pages 157–79, 2006.

Paolo Ciancarini.
Coordination models and languages as software integrators.
ACM Computing Surveys, 28(2):300–302, June 1996.

Cristiano Castelfranchi, Giovanni Pezzullo, and Luca Tummolini.
Behavioral implicit communication (BIC): Communicating with smart environments via our practical behavior and its traces.
International Journal of Ambient Computing and Intelligence, 2(1):1–12, January–March 2010.

References II

Paolo Ciancarini, Robert Tolksdorf, and Franco Zambonelli.
A survey of coordination middleware for xml-centric applications.
The Knowledge Engineering Review, 17:389–405, 12 2002.

Alessandra Di Pierro, Chris Hankin, and Herbert Wiklicky.
Probabilistic Linda-based coordination languages.
In Frank S. de Boer, Marcello M. Bonsangue, Susanne Graf, and Willem-Paul de Roever, editors, *3rd International Conference on Formal Methods for Components and Objects (FMCO'04)*, volume 3657 of *LNCS*, pages 120–140. Springer, Berlin, Heidelberg, 2005.

Tom De Wolf and Tom Holvoet.
Design patterns for decentralised coordination in self-organising emergent systems.
In *Engineering Self-Organising Systems*, pages 28–49. Springer, 2007.

Jose Luis Fernandez-Marquez, Giovanna Di Marzo Serugendo, and Josep Lluis Arcos.
Infrastructureless spatial storage algorithms.
ACM Transactions on Autonomous and Adaptive Systems (TAAS), 6(2):15, 2011.

JoseLuis Fernandez-Marquez, Giovanna Marzo Serugendo, Sara Montagna, Mirko Viroli, and JosepLluis Arcos.
Description and composition of bio-inspired design patterns: a complete overview.
Natural Computing, pages 1–25, 2012.

References III

Jose Luis Fernandez-Marquez, Giovanna Di Marzo Serugendo, and Sara Montagna.
Bio-core: Bio-inspired self-organising mechanisms core.
In *Bio-Inspired Models of Networks, Information, and Computing Systems*, volume 103 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 59–72. Springer Berlin Heidelberg, 2012.

David Gelernter.
Generative communication in Linda.
ACM Transactions on Programming Languages and Systems, 7(1):80–112, January 1985.

Luca Gardelli, Mirko Viroli, Matteo Casadei, and Andrea Omicini.
Designing self-organising MAS environments: The collective sort case.
In Danny Weyns, H. Van Dyke Parunak, and Fabien Michel, editors, *Environments for MultiAgent Systems III*, volume 4389 of *LNAI*, pages 254–271. Springer, May 2007. 3rd International Workshop (E4MAS 2006), Hakodate, Japan, 8 May 2006. Selected Revised and Invited Papers.

References IV

Stefano Mariani.
Parameter engineering vs. parameter tuning: the case of biochemical coordination in MoK.
In Matteo Baldoni, Cristina Baroglio, Federico Bergenti, and Alfredo Garro, editors, *From Objects to Agents*, volume 1099 of *CEUR Workshop Proceedings*, pages 16–23, Turin, Italy, 2–3 December 2013. Sun SITE Central Europe, RWTH Aachen University.

Stefano Mariani.
On the "local-to-global" issue in self-organisation: Chemical reactions with custom kinetic rates.
In *Eighth IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshops, SASOW 2014*, Eighth IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshops, SASOW 2014, pages 61 – 67, London, UK, September 2014. IEEE.
Best student paper award.

Thomas W. Malone and Kevin Crowston.
The interdisciplinary study of coordination.
ACM Computing Surveys, 26(1):87–119, 1994.

References V

Stefano Mariani and Andrea Omicini.
MoK: Stigmergy meets chemistry to exploit social actions for coordination purposes.
In Harko Verhagen, Pablo Noriega, Tina Balke, and Marina de Vos, editors, *Social Coordination: Principles, Artefacts and Theories (SOCIAL.PATH)*, pages 50–57, AISB Convention 2013, University of Exeter, UK, 3–5 April 2013. The Society for the Study of Artificial Intelligence and the Simulation of Behaviour.

Stefano Mariani and Andrea Omicini.
Molecules of Knowledge: Self-organisation in knowledge-intensive environments.
In Giancarlo Fortino, Costin Bădică, Michele Malgeri, and Rainer Unland, editors, *Intelligent Distributed Computing VI*, volume 446 of *Studies in Computational Intelligence*, pages 17–22. Springer, 2013.

Stefano Mariani and Andrea Omicini.
Coordination mechanisms for the modelling and simulation of stochastic systems: The case of uniform primitives.
SCS M&S Magazine, 2014.

References VI

Stefano Mariani and Andrea Omicini.
Anticipatory coordination in socio-technical knowledge-intensive environments: Behavioural implicit communication in MoK.
In Marco Gavanelli, Evelina Lamma, and Fabrizio Riguzzi, editors, *AI*IA 2015, Advances in Artificial Intelligence*, volume 9336 of *Lecture Notes in Computer Science*, chapter 8, pages 102–115. Springer International Publishing, 23–25 September 2015.
XIVth International Conference of the Italian Association for Artificial Intelligence, Ferrara, Italy, September 23–25, 2015, Proceedings.

Radhika Nagpal.
A catalog of biologically-inspired primitives for engineering self-organization.
In *Engineering Self-Organising Systems*, pages 53–62. Springer, 2004.

Andrea Omicini, Alessandro Ricci, and Mirko Viroli.
Artifacts in the A&A meta-model for multi-agent systems.
Autonomous Agents and Multi-Agent Systems, 17(3):432–456, December 2008.
Special Issue on Foundations, Advanced Topics and Industrial Perspectives of Multi-Agent Systems.

References VII

Alessandro Ricci, Andrea Omicini, Mirko Viroli, Luca Gardelli, and Enrico Oliva.
Cognitive stigmergy: Towards a framework based on agents and artifacts.
In Danny Weyns, H. Van Dyke Parunak, and Fabien Michel, editors, *Environments for MultiAgent Systems III*, volume 4389 of *LNCS*, pages 124–140. Springer, May 2007.
3rd International Workshop (E4MAS 2006), Hakodate, Japan, 8 May 2006. Selected Revised and Invited Papers.

Kjeld Schmidt and Carla Simone.
Mind the gap! Towards a unified view of CSCW.
In Rose Dieng, Alain Giboin, Laurent Karsenty, and Giorgio De Michelis, editors, *Designing Cooperative Systems: The Use of Theories and Models*, volume 58 of *Frontiers in Artificial Intelligence and Applications*, Sophia Antipolis, France, 23–26 May 2000. IOS Press.
4th International Conference on the Design of Cooperative Systems (COOP 2000), Proceedings.

References VIII

Luca Tummolini, Cristiano Castelfranchi, Alessandro Ricci, Mirko Viroli, and Andrea Omicini.
"Exhibitionists" and "voyeurs" do it better: A shared environment approach for flexible coordination with tacit messages.
In Danny Weyns, H. Van Dyke Parunak, and Fabien Michel, editors, *Environments for Multi-Agent Systems*, volume 3374 of *Lecture Notes in Artificial Intelligence*, pages 215–231. Springer, February 2005.

Akla-Esso Tchao, Matteo Risoldi, and Giovanna Di Marzo Serugendo.
Modeling self-* systems using chemically-inspired composable patterns.
In *Self-Adaptive and Self-Organizing Systems (SASO), 2011 Fifth IEEE International Conference on*, pages 109–118, October 2011.

Alan Mathison Turing.
Systems of logic based on ordinals.
Proceedings of the London Mathematical Society, 2(1):161–228, 1939.

References IX

Mirko Viroli and Matteo Casadei.
Biochemical tuple spaces for self-organising coordination.
In John Field and Vasco T. Vasconcelos, editors, *Coordination Languages and Models*, volume 5521 of *Lecture Notes in Computer Science*, pages 143–162. Springer, Lisbon, Portugal, June 2009.

Mirko Viroli, Matteo Casadei, Sara Montagna, and Franco Zambonelli.
Spatial coordination of pervasive services through chemical-inspired tuple spaces.
ACM Transactions on Autonomous and Adaptive Systems, 6(2):14:1–14:24, June 2011.

Peter Wegner.
Why interaction is more powerful than algorithms.
Communications of the ACM, 40(5):80–91, May 1997.

Brian Whitworth.
Socio-technical systems.
Encyclopedia of human computer interaction, pages 533–541, 2006.

References X

Franco Zambonelli, Gabriella Castelli, Laura Ferrari, Marco Mamei, Alberto Rosi, Giovanna Di Marzo, Matteo Risoldi, Akla-Esso Tchao, Simon Dobson, Graeme Stevenson, Yuan Ye, Elena Nardini, Andrea Omicini, Sara Montagna, Mirko Viroli, Alois Ferscha, Sascha Maschek, and Bernhard Wally.
Self-aware pervasive service ecosystems.
Procedia Computer Science, 7:197–199, December 2011.

Franco Zambonelli, Andrea Omicini, Bernhard Anzengruber, Gabriella Castelli, Francesco L. DeAngelis, Giovanna Di Marzo Serugendo, Simon Dobson, Jose Luis Fernandez-Marquez, Alois Ferscha, Marco Mamei, Stefano Mariani, Ambra Molesini, Sara Montagna, Jussi Nieminen, Danilo Pianini, Matteo Risoldi, Alberto Rosi, Graeme Stevenson, Mirko Viroli, and Juan Ye.
Developing pervasive multi-agent systems with nature-inspired coordination.
Pervasive and Mobile Computing, 17:236–252, February 2015.
Special Issue "10 years of Pervasive Computing" In Honor of Chatschik Bisdikian.

Self-organisation of Knowledge in $\mathcal{M}$olecules of $\mathcal{K}$nowledge

Stefano Mariani

DISMI
Università degli Studi di Modena e Reggio Emilia

Seminar
Autonomous Systems 16/17
June 7th, 2017