$\mathcal{M}$olecules *of* $\mathcal{K}$nowledge

Self-Organisation in Knowledge-Intensive Environments

Stefano Mariani

s.mariani@unibo.it

DISI

Alma Mater Studiorum—Università di Bologna

Seminar @ Distributed Systems Group

TU Wien, Austria - 29th April 2014

1. Context, Motivations & Goals
2. $\mathcal{M}$olecules of $\mathcal{K}$nowledge ($\mathcal{MOK}$)
   - Vision
   - Ingredients
   - Model
3. Evaluating $\mathcal{MOK}$
   - Simulations
   - Case Study
4. Conclusions & Open Questions

Outline

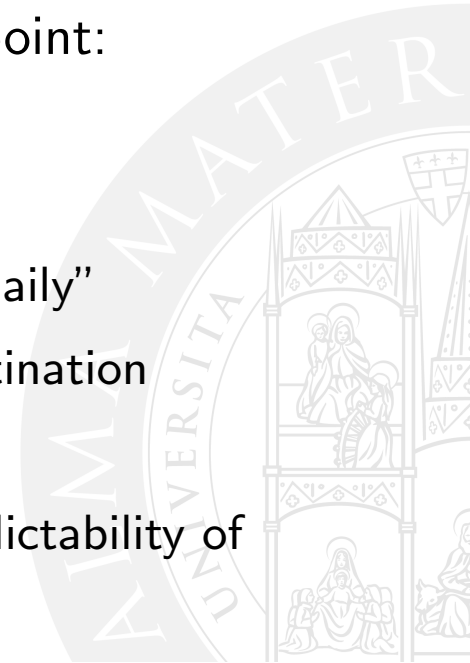
Context

Knowledge-Intensive Environments

Knowledge-Intensive Environments (KIE) [Bhatt, 2001]: "systems" combining business processes, technologies and people's skills to store, handle, make accessible – in one word, *manage* – very large repositories of (typically unstructured) information — e.g. blogs, wiki pages, online press, research portals.

Peculiar challenges from the software engineer/developer standpoint:

- **data size** — from GBs to TBs
- **scale** — from organization-wide to world-wide
- **dynamism** — new information produced/consumed "less-than-daily"
- **diversity** — both in information representation and usage destination
- **openness** — new users can enter/leave the system at any time
- **unpredictability** — KIE are often *socio-technical systems*, thus predictability of (human) behaviour is unreliable



Motivations

KIE challenges usually faced using *brute force* approaches relying on ever-increasing (hopefully, endless) (i) computational power and (ii) storage

- "big data" techniques, non-relational large-scale DBs, "data-in-the-cloud" paradigm, other buzzwords ;)

! This won't scale forever: we need alternative approaches, possibly before reaching the Moore's law upper bound and the next software crisis

"Dead data"

One possible research line departs from the following question:

why we stick to view data as a passive, "dead" thing to run algorithms upon in the traditional I/O paradigm?

Goals

The $\mathcal{M}$olecules *of* $\mathcal{K}$nowledge approach

Data is alive, information is a living thing *continuously* and *spontaneously* interacting with other information as well as with its *prosumers*, evolving itself according to such interactions [Mariani, 2011].

$\mathcal{M}$olecules *of* $\mathcal{K}$nowledge ($\mathcal{MoK}$) [Mariani and Omicini, 2013b] promotes this interpretation by relying on the following features to tackle KIE challenges:

opportunistic discovery — "trial-and-error" rather than brute-force search

locality — partial information, *decentralized* algorithms, local interactions

probability — don't strive to predict the unpredictable: rely on probability

situatedness — don't try to account for every possible scenario: be ready to adapt to the current situation

$\mathcal{MoK}$ leads to self-organising knowledge management, making "knowledge structures" appear *by emergence* from local interactions among "live" data chunks [Mariani and Omicini, 2012].

Outline

Outline

Envisioning $\mathcal{MoK}$ I

Definition

$\mathcal{M}$olecules of $\mathcal{K}$nowledge is a coordination model promoting self-organisation of knowledge in multi-agent systems (MAS), toward the idea of "self-organising workspaces" [Omicini, 2011].

Main goals:

- *autonomously* aggregate data to build more "complex" heaps of information — possibly conveying *novel* knowledge previously unknown or hidden
- *autonomously* spread such information toward potentially interested knowledge prosumers — rather than be searched *proactively*

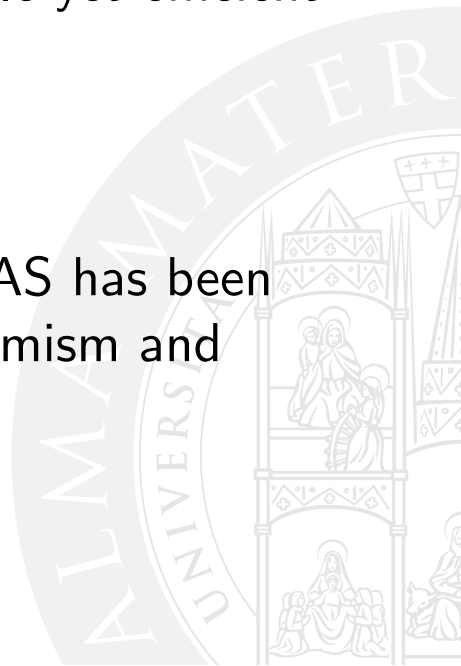
Main "sources of inspiration":

biochemical coordination — in particular, Biochemical Tuple Spaces [Viroli and Casadei, 2009]

stigmergic coordination — in particular, Behavioural Implicit Communication [Castelfranchi et al., 2010]

Envisioning MoK II

- Why a coordination model to manage data?
  - → if data is alive, then we need to properly *coordinate* such "living" data chunks' interactions [Omicini and Viroli, 2011]
- Why biochemical coordination?
  - → the *chemical metaphor* has been already shown to effectively deal with scale, openness and data size issues in MAS in a simple yet efficient way, by leveraging *locality* and *probability* features [Viroli and Casadei, 2009, Zambonelli et al., 2011]
- Why stigmergic coordination?
  - → the power of *environment-mediated* interactions in MAS has been already shown to successfully deal with diversity, dynamism and unpredictability, by leveraging on *situatedness* [Weyns et al., 2007, Castelfranchi et al., 2010]



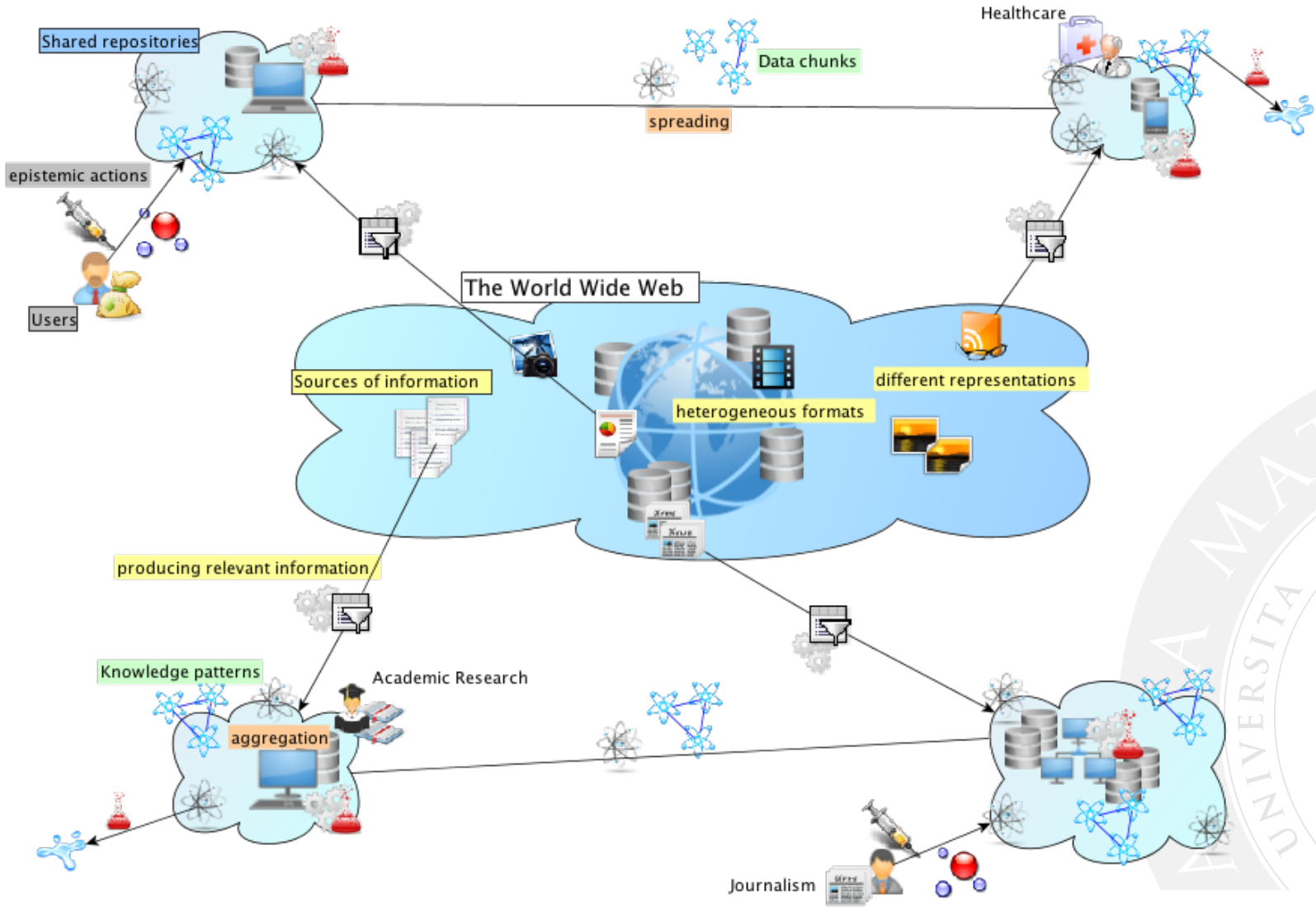
Envisioning $\mathcal{MoK}$ III

- A $\mathcal{MoK}$ system should be seen as a network of shared information repositories, in which some source entities continuously and spontaneously put data chunks
- Such data may then (i) aggregate so as to reify some (potentially) relevant "knowledge-related patterns"
  - e.g. linking two news stories talking about the same person or written by the same author

  (ii) diffuse among these networked shared spaces toward the (potentially) interested users
  - e.g. papers about elasticity should strive to reach elasticity researchers' repositories
- Users can interact with the system through epistemic actions
  - e.g. read a post, contribute to a wiki, highlight words in an article, ...

  which are tracked and exploited by the $\mathcal{MoK}$ system to influence knowledge evolution transparently to the user
  - e.g., a user highlighting a given word may imply such user being highly interested in such topics, thus $\mathcal{MoK}$ can react by, e.g., increase rank position of related topics in a search query

A MoK System I



Outline

Biochemical Coordination I

- The key idea is to coordinate any MAS entity (agents, services, data, resources) as molecules "floating" in a network of chemical compartments
- Each compartment resembles a *chemical solution*, whose "evolution" is driven by chemical reactions consuming and producing molecules — possibly from/to neighbouring compartments
- As in chemistry many reactions can occur concurrently, system evolution is driven by *probabilistic race conditions* among reactions' execution *rates*, so that certain ones are *stochastically* selected over others — usually according to the *law of mass action*[1], as in chemistry actually is [Gillespie, 1977]

[1] The rate (r_f) of a reaction is proportional (k_f) to the product of the concentrations (relative quantity) of the participating molecules ($[M'], [M'']$): $r_f = k_f[M'][M'']$, where k_f is the *rate constant* and, in chemistry, is a function of molecules "affinity".

Biochemical Coordination II

Biochemical Tuple Spaces

Biochemical Tuple Spaces (BTS) [Viroli and Casadei, 2009] implement biochemical coordination upon a *tuple space -based* setting.

- → molecules are reified into tuples, which are equipped with an activity/pertinency value — roughly resembling chemical concentration[2]
- → chemical reactions are reified into coordination laws, evolving tuples' concentration and possibly diffusing them to neighboring tuple spaces
- → chemical compartments are reified into tuple spaces, acting as *chemical solutions simulators* — that is, they execute the Gillespie algorithm [Gillespie, 1977] and carry out diffusion

[2] In BTS, the concentration is the multiplicity of the tuple in the tuple space, not is relative quantity w.r.t. other tuples in the same space.

Biochemical Coordination III

$\mathcal{MoK}$ & BTS

$\mathcal{MoK}$ is roughly based on the BTS model, but, with a few fundamental differences.

- ! $\mathcal{MoK}$ is focussed on information management solely
- ! $\mathcal{MoK}$ reactions execution does not strictly follow the law of mass action[3]
- ! $\mathcal{MoK}$ has a fixed set of reactions, as opposed to BTS in which they have to be programmed according to the application at hand — this allows to study self-organising properties holding for all $\mathcal{MoK}$-based systems
- ! $\mathcal{MoK}$ adds to the BTS model concepts and mechanisms borrowed from *stigmergic coordination* — see following slides

[3] Implications of this aspect are too specific to be worth being described here. The interested reader is referred to [Mariani, 2013].

Behavioural Implicit Communication I

In social systems, interactions between individuals are usually *mediated by the environment*, which "records" the traces (potentially) left by individuals' actions — the same holds for artificial (MAS) and hybrid (socio-technical) systems [Weyns et al., 2007].

Stigmergy

Trace-based communication is at the root of the notion of stigmergy, firstly introduced in the biological study of social insects [Grassé, 1959].

Behavioural Implicit Communication

Behavioural Implicit Communication (BIC) [Castelfranchi et al., 2010] generalises the notion of stigmergy by taking into account also actions "as a whole" (not only their traces) which can be *observed* and *interpreted* so as to promote coordination.

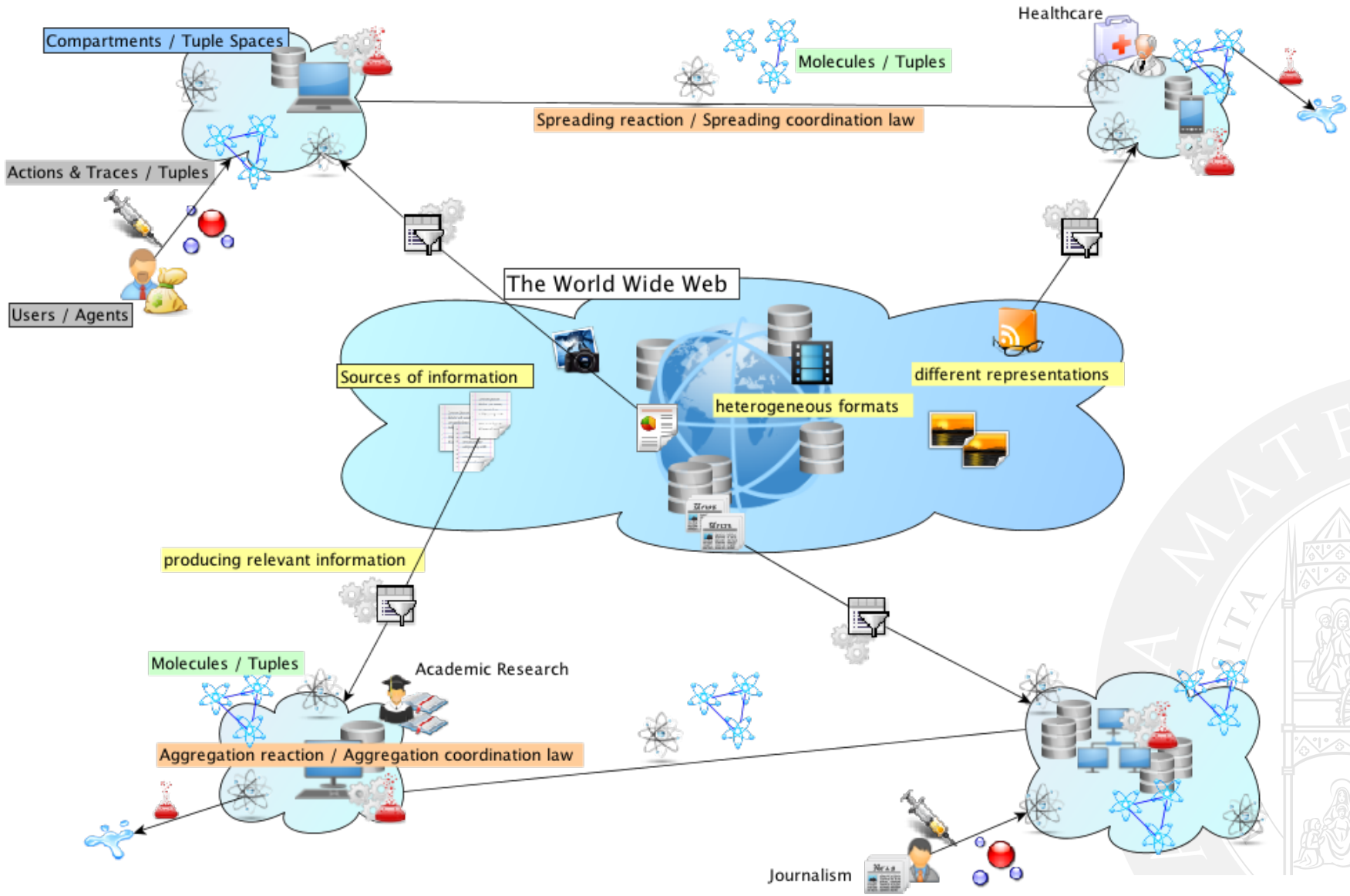
Behavioural Implicit Communication II

$\mathcal{MOK}$ & BIC

$\mathcal{MOK}$ relies on the concepts of observable action (from BIC) and of trace (from stigmergy) to make users' actions influence system behaviour [Mariani and Omicini, 2013a].

- In $\mathcal{MOK}$, any user action leaves traces (in the form of tuples) in the environment — the tuple space acting as information repository
- Such actions and their traces are then used by the $\mathcal{MOK}$ system itself – by $\mathcal{MOK}$ coordination laws – to drive evolution of information — creating tuples, increasing their concentration, moving them between tuple spaces

A MoK System II



Outline

Overview

The $\mathcal{M}$olecules of $\mathcal{K}$nowledge model features the following abstractions [Mariani and Omicini, 2013b]:

Seeds — The "sources" of information

Compartments — The information repositories all other $\mathcal{MOK}$ abstractions (conceptually) belong to

Catalysts — "Prosumers" of information (both producer and consumer users)

Atoms — The "primitive" unit of information in $\mathcal{MOK}$ (always produced by a seed)

Molecules — The "composite" unit of information in $\mathcal{MOK}$ (interaction "patterns" found by $\mathcal{MOK}$ itself among different atoms are reified into molecules)

Enzymes — The reification of catalysts' epistemic actions

Reactions — The "laws of nature" driving $\mathcal{MOK}$ compartments evolution

Seeds

Definition

MoK seeds are the sources of information, *continuously* and *spontaneously* injecting data pieces (atoms) into the workspace (compartment) they belong to.

→ they should represent information sources in a *complete* and *machine-* as well as *human-readable* way, so that any information chunk could be both manually and automatically extracted from them and used at any time

- e.g. an XML-tagged document along with its XML Schema

→ once information has been extracted, it has to be packed into atoms – either automatically or manually – and *perpetually* injected into the compartment according to a certain rate — which may vary as time passes

- e.g. an XML parser can read the document, extract the tagged items, wrap them into atoms then put them in the repository every 30 minutes for a week

Compartments

Definition

$\mathcal{MoK}$ compartments are the *active* repositories of information storing seeds, atoms, molecules and enzymes, also responsible for scheduling and executing $\mathcal{MoK}$ reactions.

- → they are the "computational *loci*" in which $\mathcal{MoK}$ reactions take place consuming and producing $\mathcal{MoK}$ atoms, molecules and enzymes so as to resemble *chemical solutions dynamics*
  - e.g. a $\mathcal{MoK}$ reaction could aggregate two atoms into a molecule if they come from the same seed
- → also, they are the "topological abstraction" giving $\mathcal{MoK}$ the notions of *locality* and *neighborhood* and enabling diffusion of atoms and molecules to take place
  - e.g. a given $\mathcal{MoK}$ compartment may be connected only to a subset of the whole network of compartments in a certain $\mathcal{MoK}$ system, thus defining a precise neighbourhood atoms and molecules can spread to

Catalysts

Definition

$\mathcal{MoK}$ catalysts are the users of the $\mathcal{MoK}$ system, the *prosumers* both exploiting and influencing $\mathcal{MoK}$ self-organisation services.

→ *exploiting*, because they are typically interested in storing, manipulating and retrieving the information they need to carry out their business

- e.g. journalists could search for press articles, highlight some relevant phrase, store it for later usage, then rearrange such news pieces in a novel story

→ *influencing*, because their epistemic actions are properly reified, observed and used by the $\mathcal{MoK}$ system to autonomously and spontaneously evolve such information

- e.g. the keywords in the search query may be used to increase concentration of related molecules, atoms corresponding to highlighted phrases may be aggregated into a single molecule, the novel story may be reified as a new seed

Atoms

Definition

Atoms are the smallest, "atomic" unit of information in $\mathcal{MoK}$. They are *continuously* and *spontaneously* injected by a seed into the compartment it belongs to.

$$\texttt{atom}(src, val, attr)_c$$

- src — is the seed the atom is born from (either its reference or its unique identifier)
- val — is the content of the atom, the information element it stores (e.g. a single word, a phrase, an entire paper, ...)
- attr — is any kind of metadata defining the content, meant to help *semantic* interpretation (e.g. the ontology concept defining the content)
- c — is the concentration of the atom (any form of concentration, e.g. BTS multiplicity or Gillespie's chemical concentration)

Molecules

Definition

Molecules are the evolving, composite unit of information in $\mathcal{MOK}$. They are *continuously* and *spontaneously* produced by the compartment they belong to so as to reify some meaningful knowledge.

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

Atoms — is the set of atoms[4] aggregated into the molecule (e.g. atoms coming from the same seed, with *semantically related* content, with "somehow related" metadata — see slide 34)

c — is the concentration of the molecule

[4] Could be the actual atoms, a reference to them, their unique identifier, ...

Enzymes

Definition

Enzymes are the reification of catalysts' epistemic actions over the information stored within their compartment. They are *automatically* produced by the compartment to be exploited into $\mathcal{MoK}$ reactions.

$$\texttt{enzyme}(\mathit{Molecule}, f)_c$$

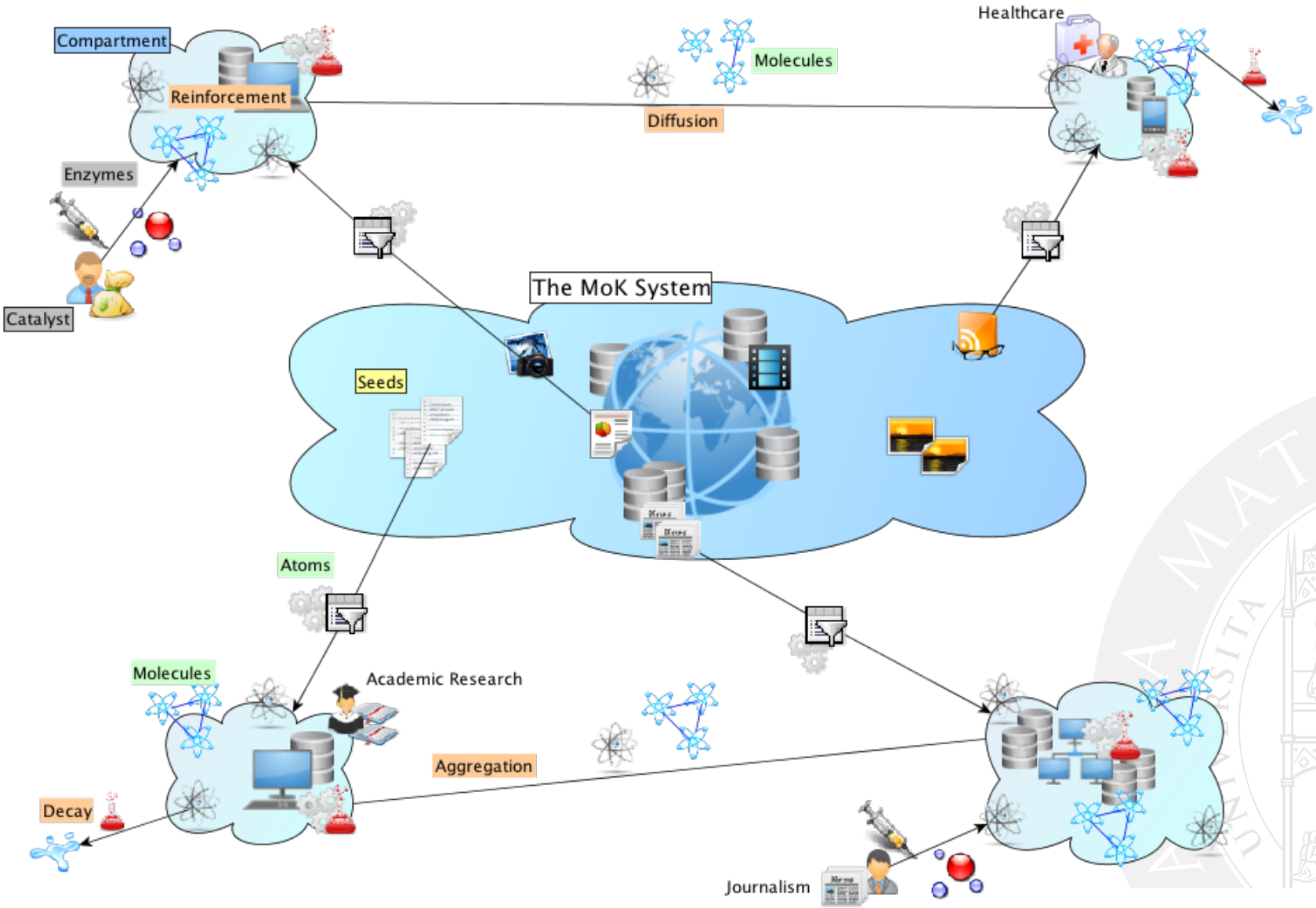
Molecule — is the molecule[5] subject of the action reified by the enzyme (e.g. when catalyst C performs the "highlighting action" on molecule M, an $\texttt{enzyme}(M)_c$ is produced)

f — is the entity of the concentration increment

c — is the concentration of the enzyme

[5] Enzymes apply also to atoms. To ease notation, atoms can be regard as molecules with a single element. Thus, in the following, the term molecule can indicate also atoms.

A MoK System III



Reactions I

Definition

Reactions are the *general purpose coordination laws* resembling biochemical reactions in $\mathcal{MoK}$. They drive information evolution within the compartment they are "installed" in by governing molecules interactions.

Four reactions are considered so far[6] in the $\mathcal{MoK}$ model:

Aggregation — Bounds together semantically related molecules

Diffusion — Move molecules between neighbour compartments

Reinforcement — Consume enzymes to increase the concentration of their molecule

Decay — Decrease concentration of molecules as time passes

[6] This is by no means the definite set of available reactions. Work on studying the expressiveness of this "core" reactions set is in progress.

Reactions II

Aggregation

Aggregation reaction consumes N molecules[a] to produce a single molecule which aggregates the consumed ones based on some *semantic similarity measure* — see slide 34.

[a] Whenever concentration c is not specified, it is assumed to be 1.

$$\texttt{molecule}(Atoms^1) + \ldots + \texttt{molecule}(Atoms^n) \xrightarrow{r_{agg}} \texttt{molecule}(Atoms^1 \uplus \ldots \uplus Atoms^n) + Residual(Atoms^1 \div \ldots \div Atoms^n)$$

- $\uplus$ — denotes union of semantically related atoms
- $\div$ — denotes union of unrelated atoms
- *Residual* — denotes the set of unrelated atoms to be put back in the compartment as distinct items

Reactions III

Diffusion

Diffusion reaction moves a molecule form a source compartment to a destination one. The source compartment is where the molecule currently is, the destination one is *uniformly* chosen among those in its *neighbourhood*.

$$\{Molecule \cup Molecules\}_{\sigma^1} + \{Molecules\}_{\sigma^2} \xrightarrow{r_{diff}} \{Molecules\}_{\sigma^1} + \{Molecules \cup Molecule\}_{\sigma^2}$$

$\{\cdot\}_{\sigma^i}$ — denotes the molecules within the compartment identified by σ^i

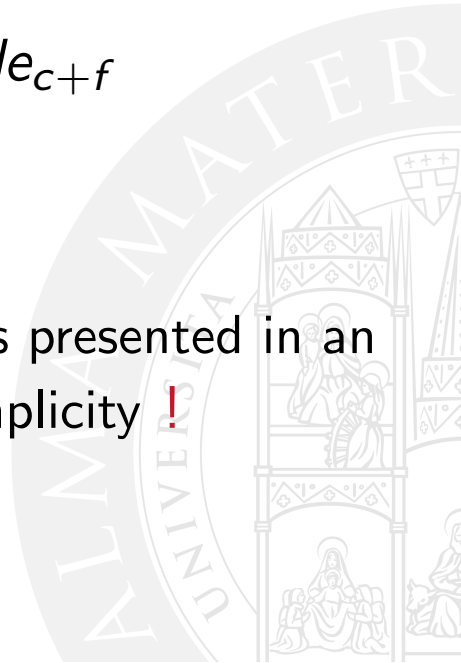
Reactions IV

Reinforcement

Reinforcement reaction consumes an enzyme to increase the concentration c of the molecule it is related to by f.

$$\texttt{enzyme}(Molecule, f) + Molecule_c \xrightarrow{r_{reinf}} Molecule_{c+f}$$

! In [Mariani and Omicini, 2013a], the reinforcement reaction is presented in an enhanced version, not reported here for the sake of simplicity !



Reactions V

Decay

Decay reaction decreases the concentration c of a uniformly chosen molecule within the compartment by 1. If $c < 1 \rightarrow c = 0$.

$$Molecule_c \xrightarrow{r_{dec}} Molecule_{c-1}$$

The Function $\mathcal{F}_{\mathcal{MoK}}$

? What is the semantic similarity measure used by aggregation reaction ?

$\mathcal{F}_{\mathcal{MoK}}$

It's the value $m \in [0, 1]$ given by the function $\mathcal{F}_{\mathcal{MoK}}$, defined as

$$\mathcal{F}_{\mathcal{MoK}}\colon\ \mathit{Molecule}^1 \times \mathit{Molecule}^2 \mapsto m \in [0, 1]$$

- Function $\mathcal{F}_{\mathcal{MoK}}$ should be designed such as m represents the extent to which $\mathit{Molecule}^1$ and $\mathit{Molecule}^2$ are semantically related
- Usually, $\mathcal{F}_{\mathcal{MoK}}$ is application-specific — e.g. LINDA [Gelernter, 1985] syntactical matching, thus $m = 0 \vee 1$, or some kind of *fuzzy*, ontology-based matching returning continuous values of $m \in [0, 1]$

! Function $\mathcal{F}_{\mathcal{MoK}}$ is used also to match $\mathcal{MoK}$ reactions' reactant *templates* against actual molecules within the compartment !

A $\mathcal{MoK}$ System IV

- A $\mathcal{MoK}$ system is a network of $\mathcal{MoK}$ compartments, in which $\mathcal{MoK}$ seeds continuously and spontaneously inject $\mathcal{MoK}$ atoms
- $\mathcal{MoK}$ atoms (and eventually $\mathcal{MoK}$ molecules) may then aggregate, diffuse, reinforce and decay, driven by $\mathcal{MoK}$ reactions and $\mathcal{MoK}$ enzymes produced by $\mathcal{MoK}$ catalysts' behaviour
- $\mathcal{MoK}$ reactions are scheduled and executed by $\mathcal{MoK}$ compartments according to Gillespie's *chemical dynamics* simulation algorithm [Gillespie, 1977], so as to promote self-organisation based on locality, situatedness and probability

  probability — selection of reactants in $\mathcal{MoK}$ reactions is probabilistically based on the $\mathcal{F}_{\mathcal{MoK}}$ function; scheduling of reactions is probabilistic according to Gillespie; execution of reactions is probabilistic in time taken

  situatedness — decay and diffusion reactions enforce, respectively, situatedness in time and space; reinforcement supports situatedness "in context"

  locality — compartments execute $\mathcal{MoK}$ reactions locally; diffusion is based on neighbourhood

Outline

$\mathcal{MoK}$ Early Stage Evaluation

- To date, $\mathcal{MoK}$ is still little more than an (hopefully, interesting) idea: no large-scale $\mathcal{MoK}$ systems have been deployed in the real world, no large-scale simulations of all $\mathcal{MoK}$ features have been carried out
- Nevertheless, a prototype implementation of a "$\mathcal{MoK}$ engine" exists, deployed upon the TuCSoN coordination infrastructure[7] [Omicini and Zambonelli, 1999]
  - Roughly speaking, TuCSoN tuple centres [Omicini and Denti, 2001] reify $\mathcal{MoK}$ compartments and a ReSpecT specification [Omicini, 2007] therein installed implements $\mathcal{MoK}$ reactions
- Also, simulations regarding $\mathcal{MoK}$ reactions interplay and effectiveness in exhibiting self-organising behaviours have been carried out
  - In particular, using the Bio-PEPA framework [Ciocchetta and Hillston, 2009] to face the problem of *parameter engineering* in $\mathcal{MoK}$ reactions' rate expressions [Mariani, 2013]

[7] Home page at `http://tucson.unibo.it`

Outline

Bio-PEPA & $\mathcal{MoK}$ I

Bio-PEPA

Bio-PEPA [Ciocchetta and Hillston, 2009] is a framework (hence language + Eclipse plugin) to model, simulate and monitor biochemical processes.

Main features:

- custom *kinetic laws* (rate expressions) represented as functional expressions
- definition of *stoichiometry* and role played by the species (reactant, product, enzyme, etc.) in a given reaction
- semantics is formally well-founded on *CTMC semantics*

We used Bio-PEPA to simulate $\mathcal{MoK}$ reactions "behaviour" if they were scheduled according to different rate expressions.

Goal

"Best" rates $\rightarrow$ "best" $\mathcal{MoK}$ reactions $\rightarrow$ "best" self-organising behaviour.

$\mathcal{MoK}$ Reactions Behaviour I

$\mathcal{MoK}$ Decay

$$r_{dec} = log(time) + \frac{[Seed]}{[Molecule]}$$

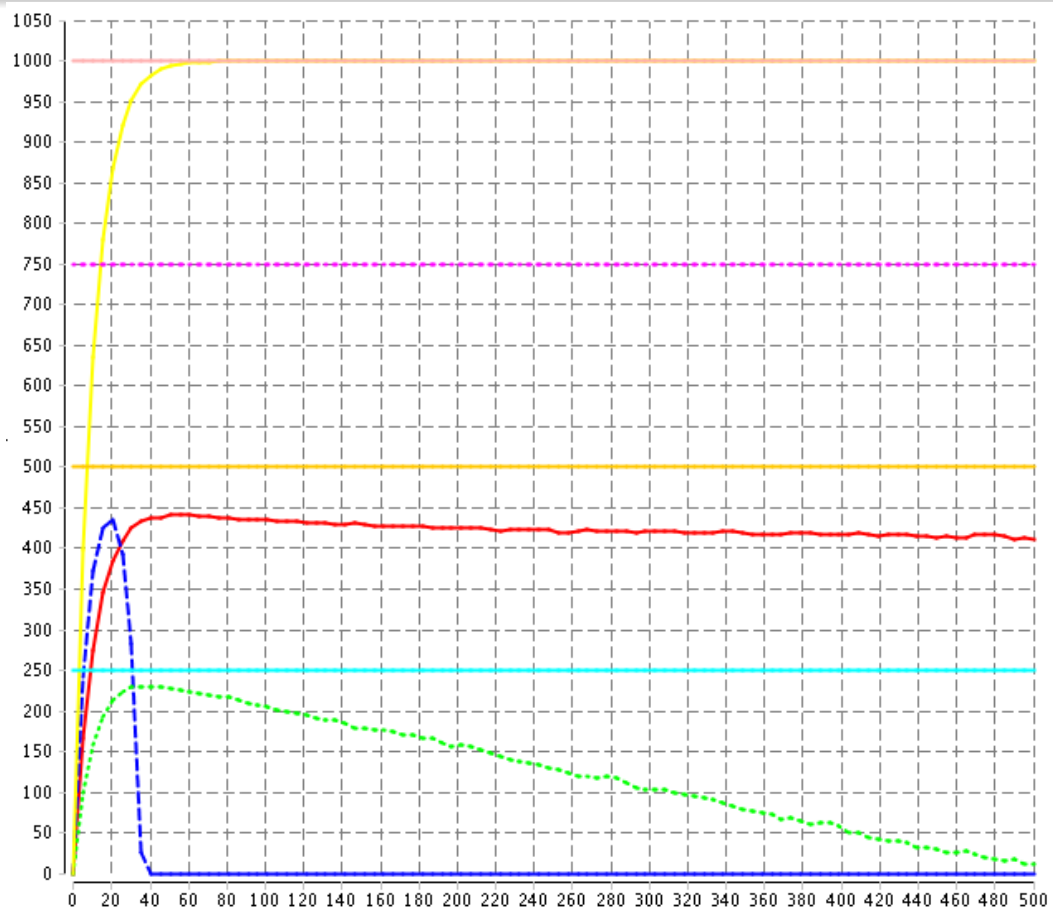


Figure : $\mathcal{MoK}$ decay shown by red line. Orange line above is "saturation level" [Mariani, 2013].

$\mathcal{MoK}$ Reactions Behaviour II

$\mathcal{MoK}$ Reinforcement

$$r_{reinf} = \frac{[Molecule]}{[Seed]} \text{ (time window} = t_{300} \rightarrow t_{700})$$

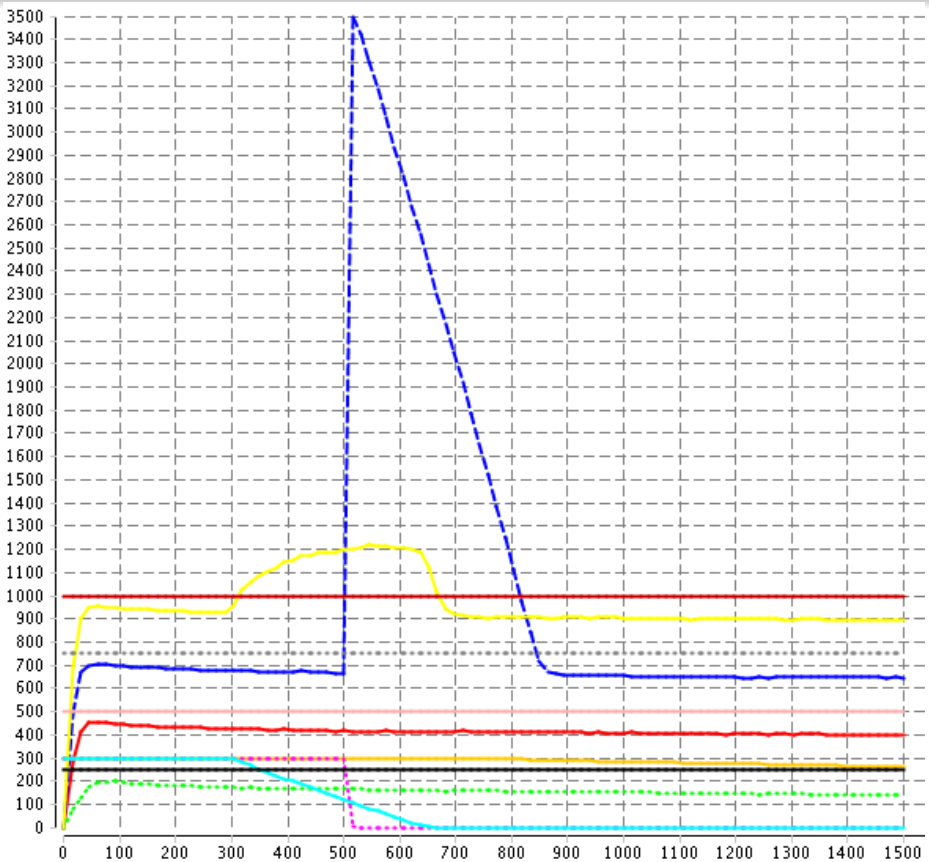


Figure : $\mathcal{MoK}$ reinforcement shown by yellow line. Red line traversed is saturation level.

$\mathcal{MoK}$ Reactions Behaviour III

$\mathcal{MoK}$ Diffusion

$$r_{diff} = \frac{\{[Molecule]\}_{\sigma_s}}{\{[Molecule]\}_{\sigma_d}} + K \ (K = parameter)$$

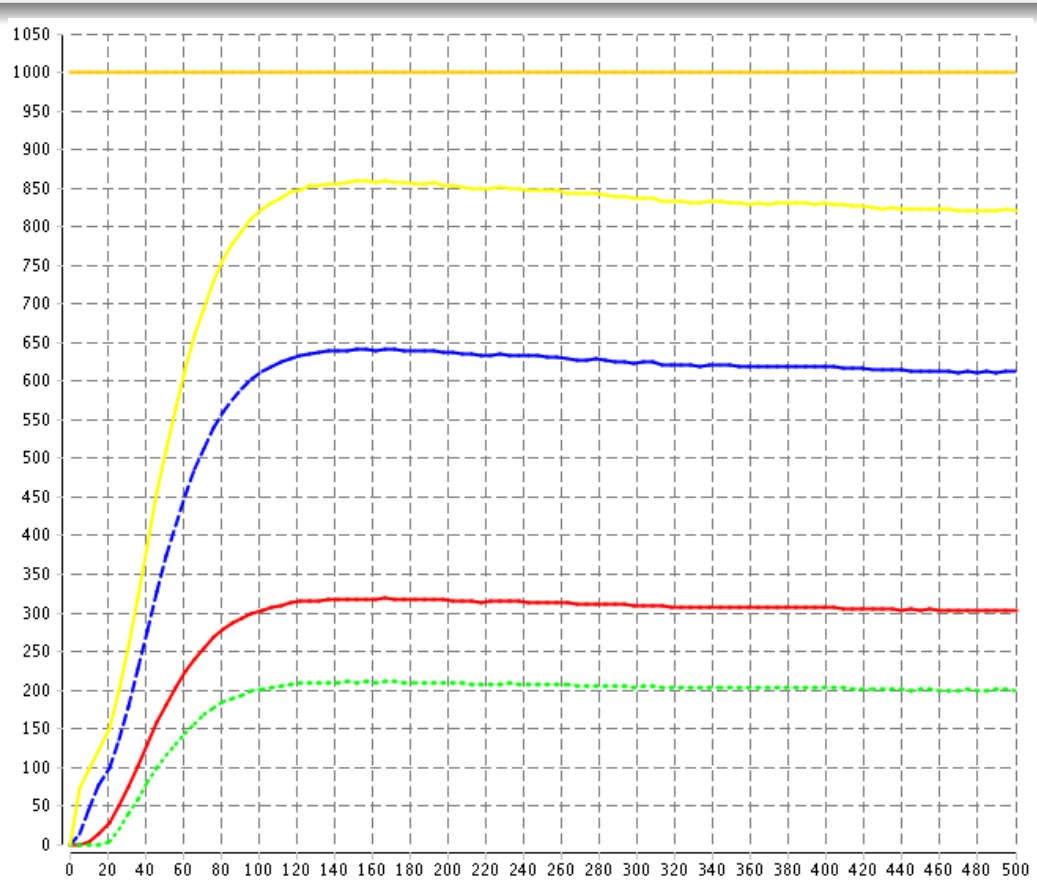


Figure : $\mathcal{MoK}$ diffusion shown by blue, red, green lines. Yellow line above is $\{[Molecule]\}_{\sigma_s}$.

$\mathcal{MoK}$ Reactions Behaviour IV

Figure : $\mathcal{MoK}$ reactions interplay.

$\mathcal{MoK}$ Model Refined

Before

$$\{Molecule \cup Molecules\}_{\sigma^1} + \{Molecules\}_{\sigma^2} \xrightarrow{r_{diff}} \{Molecules\}_{\sigma^1} + \{Molecules \cup Molecule\}_{\sigma^2}$$

$$\texttt{enzyme}(Molecule, f) + Molecule_c \xrightarrow{r_{reinf}} Molecule_{c+f}$$

$$Molecule_c \xrightarrow{r_{dec}} Molecule_{c-1}$$

$\Rightarrow$

$\Rightarrow$

$\Rightarrow$

After

$$\{Molecule \cup Molecules\}_{\sigma^1} + \{Molecules\}_{\sigma^2} \xrightarrow{\frac{\{[Molecule]\}\sigma_s}{\{[Molecule]\}\sigma_d} + K} \{Molecules\}_{\sigma^1} + \{Molecules \cup Molecule\}_{\sigma^2}$$

$$\texttt{enzyme}(Molecule, f) + Molecule_c \xrightarrow{\frac{[Molecule]}{[Seed]}} Molecule_{c+f}$$

$$Molecule_c \xrightarrow{log(time) + \frac{[Seed]}{[Molecule]}} Molecule_{c-1}$$

! $\mathcal{MoK}$ aggregation reaction not simulated with Bio-PEPA — see [Mariani, 2013] !

! In [Mariani, 2013] also an injection reaction is considered !

Outline

The $\mathcal{MoK}$-News Scenario I

Why News?

New management systems are a prominent example of socio-technical KIE.

heterogeneity — news "sources" can be virtually *anything* — "official" press, blogs, social networks

ubiquity — netbooks, tablets and smartphones made information production, sharing and consumption as *pervasive* as never before

unpredictability — if journalists can roughly follow some *structured* workflow, the same does not hold for contributors to the so-called "citizen journalism"

$\mathcal{MoK}$ & News

New management systems are thus well suited for $\mathcal{MoK}$.

The $\mathcal{MoK}$-News Scenario II

- In [Mariani and Omicini, 2012], we took IPTC's[8] technical standards regarding news management, in particular:
  - `NewsML` — an *XML-based* tagging language meant to ease news sharing by relying on `NewsCodes` *ontologies*
  - `NITF` — an XML-based tagging language meant to enrich the news content, again by relying on `NewsCodes`
- We identified an abstract mapping between a $\mathcal{MoK}$ atom and a `NewsML/NITF` *tag*, in particular:
  - → $\texttt{atom}(src, val, \texttt{sem}(tag, catalog))_c$ where *tag* ::= `NewsML/NITF` *tag* and *catalog* ::= `NewsCodes` *uri*
- We implemented the domain-specific $\mathcal{MoK}$ model – called $\mathcal{MoK}$-News – on an existing platform for distributed coordination in MAS, in particular:
  - → on TuCSoN, by using its tuple centres as $\mathcal{MoK}$ compartments, in which a ReSpecT specification implements the Gillespie algorithm; TuCSoN tuples as $\mathcal{MoK}$ seeds/atoms/molecules/enzymes/reactions and TuCSoN agents as $\mathcal{MoK}$ catalysts

[8] `http://www.iptc.org/site/Home/About/`

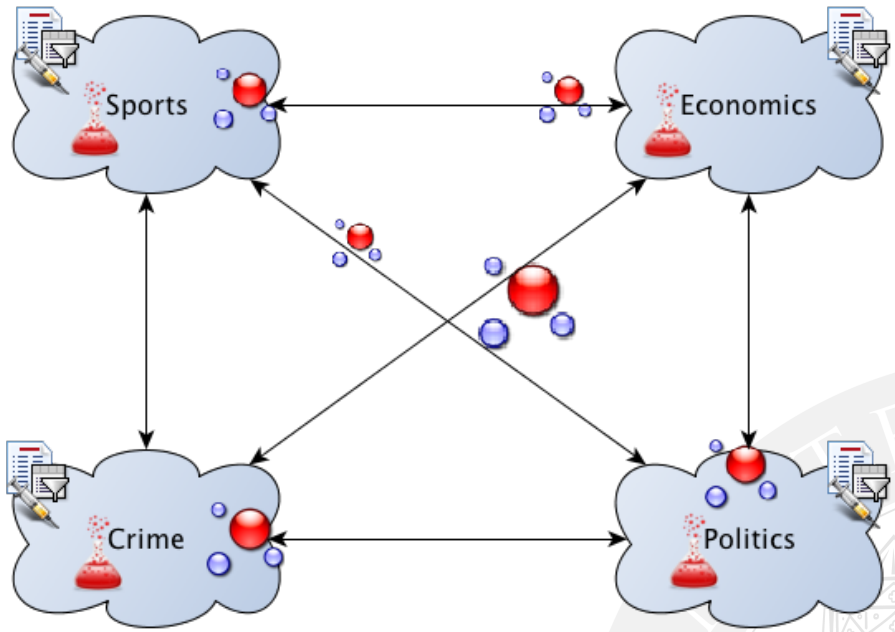
A $\mathcal{MoK}$ System V

A $\mathcal{MoK}$-`News` system could then be deployed to an online magazine newsroom as follows:

- journalists may be given a smartphone to use as their workspace, running a $\mathcal{MoK}$ compartment decorated with a suitable GUI letting them carry out their usual working habits — searching for news in the web, storing some for later use, creating and editing stories, ...
- within compartments, news sources (seeds), news pieces (atoms/molecules) and journalists' actions traces (enzymes) live and interact in a completely autonomous way so as to best support journalists' workflow — seeds increasing/decreasing atoms injection rate, molecules aggregating to automagically compose news stories, enzymes increasing/decreasing relevance (concentration) of news pieces, ...
- as they use the system, journalists implicitly drive its behaviour (e.g. reaction rates) toward their needs, *closing the feedback loop* providing $\mathcal{MoK}$-`News` with self-adaptive capabilities — combination of diffusion and reinforcement, driven by journalists' enzymes, enables smart migration

"Smart Migration" in $\mathcal{MoK}$-News I

- Imagine the $\mathcal{MoK}$-News compartments topology to the right to be deployed
- At "bootstrap", news stories are equally distributed among the compartments — since no *a priori* knowledge about journalists' interest topics is assumed
- diffusion and decay reactions *apply to all* molecules with *same rate expression*



! Eventually, the news "spatial organisation" will change according to journalists' interactions — in particular, thanks to the interplay between enzymes and $\mathcal{MoK}$ diffusion, reinforcement and decay reactions

"Smart Migration" in $\mathcal{MoK}$-News II

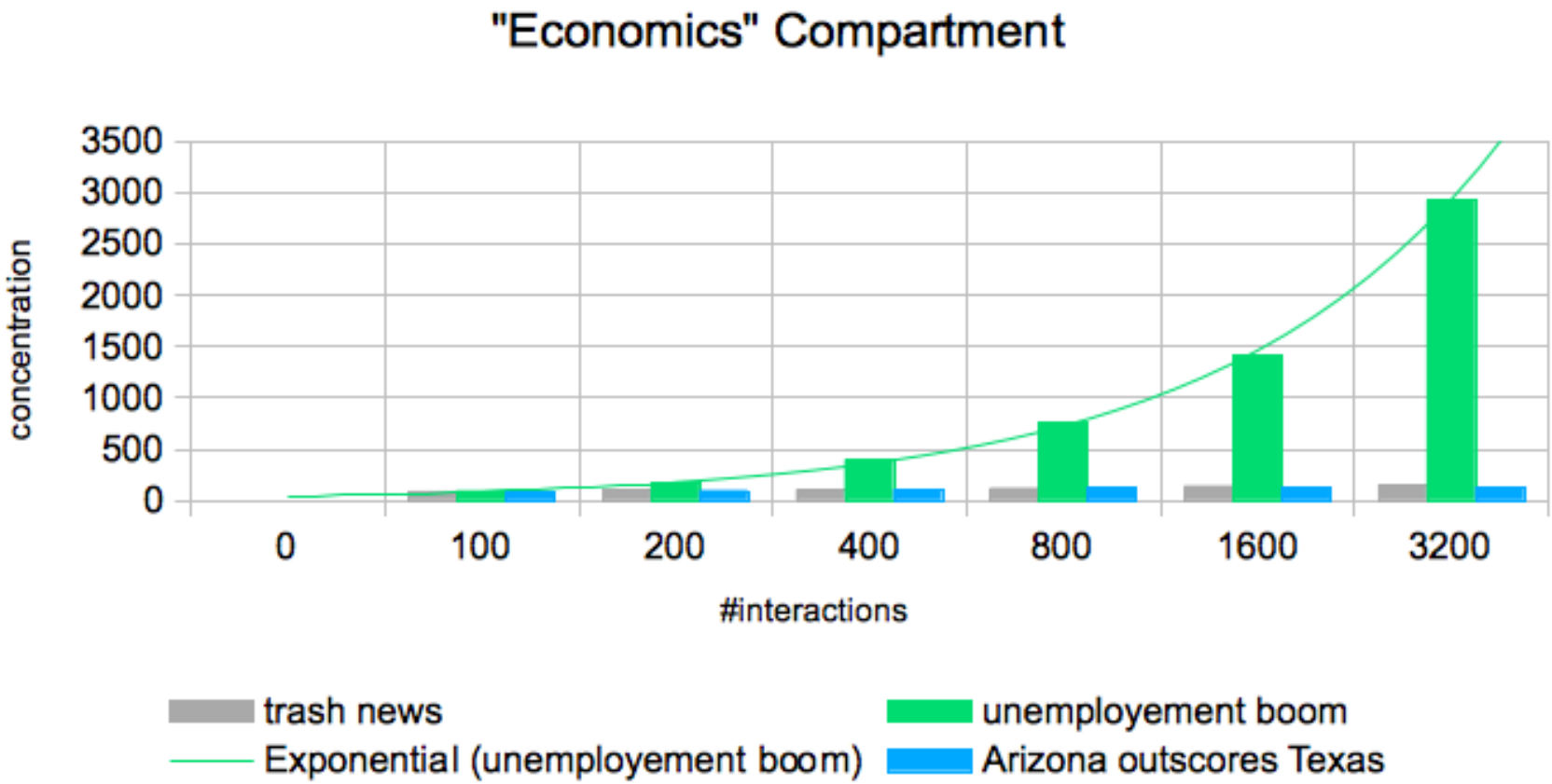


Figure : "Economics news" tend to be the most numerous (hence, relevant) in the "economics compartment" — that is, the one belonging to a journalist interested in economics.

"Smart Migration" in $\mathcal{MoK}$-News III

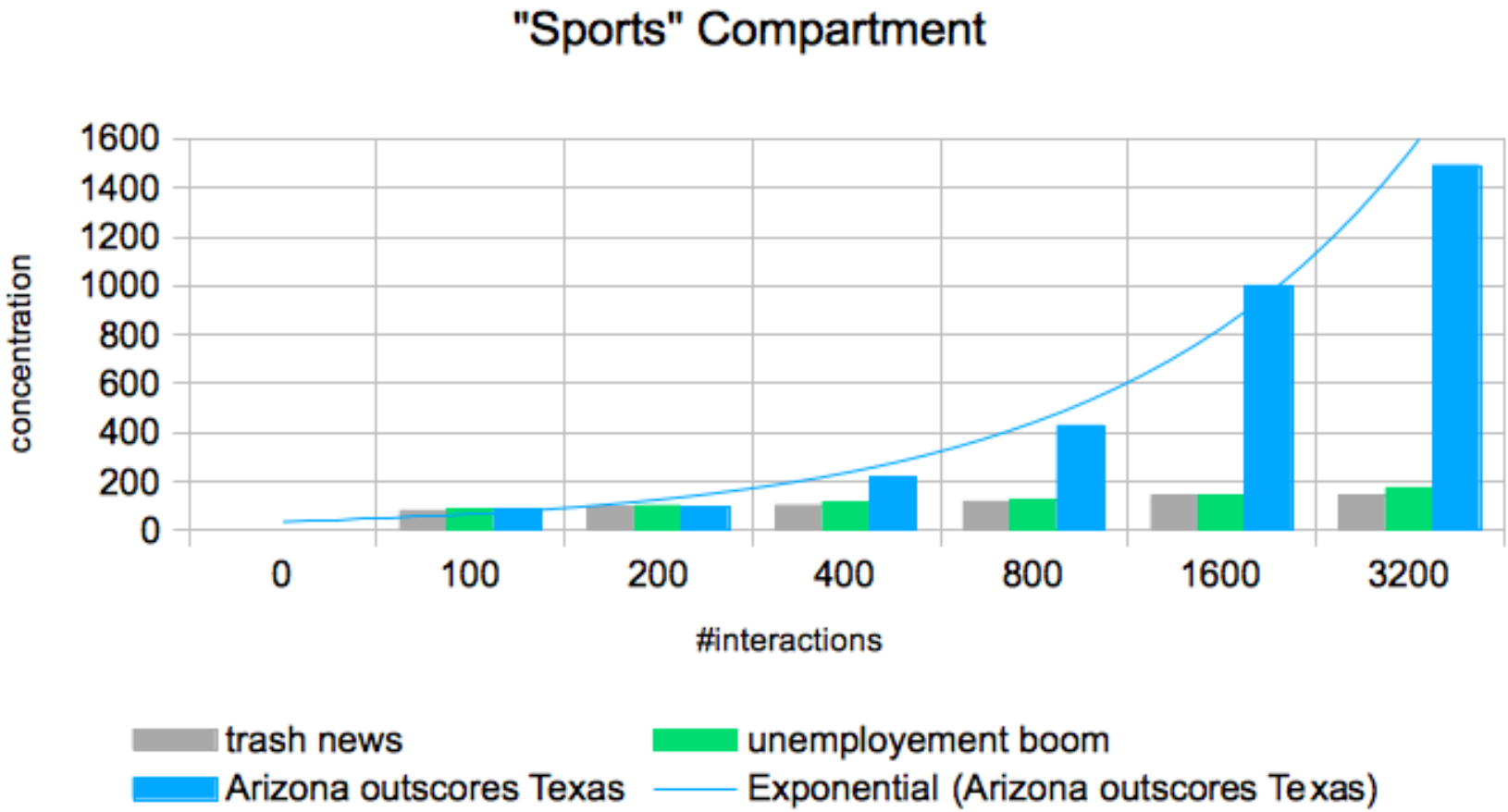
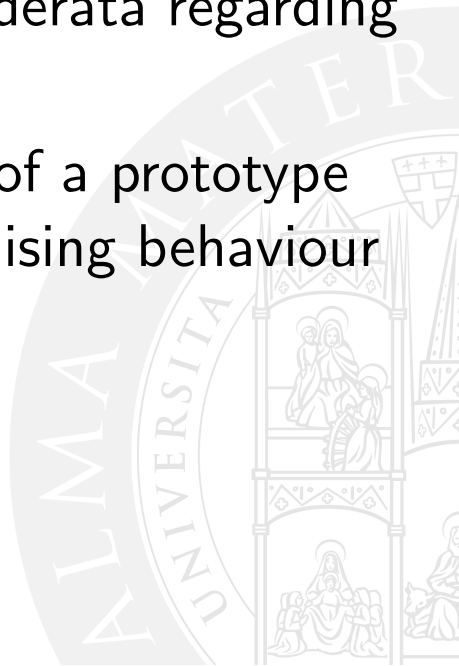


Figure : In another compartment (e.g. "sports compartment"), they are not relevant at all, because the journalist working therein is focussed on other topics (e.g. "sports news")

Outline

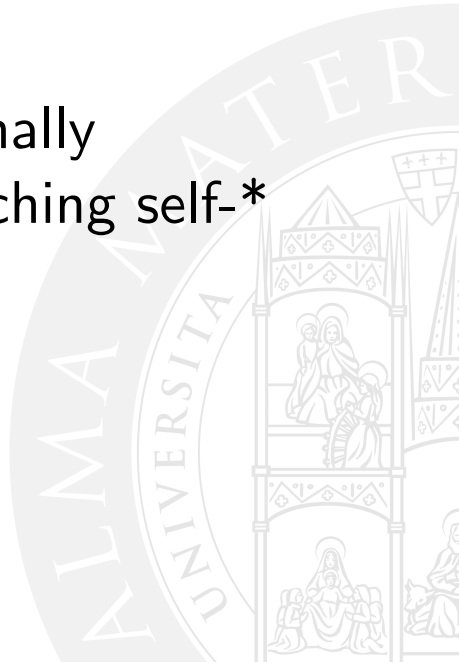
Conclusion

- The $\mathcal{MoK}$ model for self-organisation of knowledge promotes a novel interpretation of information as "living things" autonomously interacting and evolving
- Early simulations have confirmed some of $\mathcal{MoK}$'s desiderata regarding run-time behaviour of its reactions
- Early evaluation on a case study implemented on top of a prototype "$\mathcal{MoK}$ engine" also exhibited an interesting self-organising behaviour regarding spatial displacement of information



Open Questions

- Implementing a full-fledged $\mathcal{MoK}$ infrastructure is still far from current $\mathcal{MoK}$ state
- Many skills are required to do so — e.g. knowledge representation, semantic matching, machine learning, ...
- Also the $\mathcal{MoK}$ model needs to be completed and formally investigated, especially about its expressiveness in reaching self-* behaviours



References I

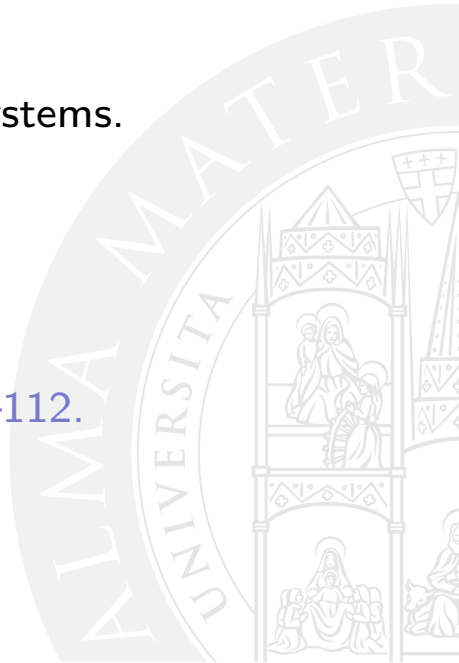
Bhatt, G. D. (2001).
Knowledge management in organizations: examining the interaction between technologies, techniques, and people.
Journal of Knowledge Management, 5(1):68–75.

Castelfranchi, C., Pezzullo, G., and Tummolini, L. (2010).
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.

Ciocchetta, F. and Hillston, J. (2009).
Bio-PEPA: A framework for the modelling and analysis of biological systems.
Theoretical Computer Science, 410(33–34):3065 – 3084.
Concurrent Systems Biology: To Nadia Busi (1968–2007).

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

Gillespie, D. T. (1977).
Exact stochastic simulation of coupled chemical reactions.
The Journal of Physical Chemistry, 81(25):2340–2361.



References II

Grassé, P.-P. (1959).
La reconstruction du nid et les coordinations interindividuelles chez Bellicositermes natalensis et Cubitermes sp. la théorie de la stigmergie: Essai d'interprétation du comportement des termites constructeurs.
Insectes Sociaux, 6(1):41–80.

Mariani, S. (2011).
Molecules of knowledge: a new approach to knowledge production, management and consumption.
Master's thesis, II Facoltà di Ingegneria Informatica – Alma Mater Studiorum Università di Bologna.

Mariani, S. (2013).
Parameter engineering vs. parameter tuning: the case of biochemical coordination in MoK.
In Baldoni, M., Baroglio, C., Bergenti, F., and Garro, A., editors, *From Objects to Agents*, volume 1099 of *CEUR Workshop Proceedings*, pages 16–23, Turin, Italy. Sun SITE Central Europe, RWTH Aachen University.
XIV Workshop (WOA 2013). Workshop Notes.

References III

Mariani, S. and Omicini, A. (2012).
Self-organising news management: The *Molecules of Knowledge* approach.
In Fernandez-Marquez, J. L., Montagna, S., Omicini, A., and Zambonelli, F., editors, *1st International Workshop on Adaptive Service Ecosystems: Natural and Socially Inspired Solutions (ASENSIS 2012)*, pages 11–16, SASO 2012, Lyon, France.
Pre-proceedings.

Mariani, S. and Omicini, A. (2013a).
MoK: Stigmergy meets chemistry to exploit social actions for coordination purposes.
In Verhagen, H., Noriega, P., Balke, T., and de Vos, M., editors, *Social Coordination: Principles, Artefacts and Theories (SOCIAL.PATH)*, pages 50–57, AISB Convention 2013, University of Exeter, UK. The Society for the Study of Artificial Intelligence and the Simulation of Behaviour.

Mariani, S. and Omicini, A. (2013b).
Molecules of Knowledge: Self-organisation in knowledge-intensive environments.
In Fortino, G., Bădică, C., Malgeri, M., and Unland, R., editors, *Intelligent Distributed Computing VI*, volume 446 of *Studies in Computational Intelligence*, pages 17–22. Springer.
6th International Symposium on Intelligent Distributed Computing (IDC 2012), Calabria, Italy, 24-26 September 2012. Proceedings.

References IV

Omicini, A. (2007).
Formal ReSpecT in the A&A perspective.
Electronic Notes in Theoretical Computer Science, 175(2):97–117.
5th International Workshop on Foundations of Coordination Languages and Software Architectures (FOCLASA'06), CONCUR'06, Bonn, Germany, 31 August 2006. Post-proceedings.

Omicini, A. (2011).
Self-organising knowledge-intensive workspaces.
In Ferscha, A., editor, *Pervasive Adaptation. The Next Generation Pervasive Computing Research Agenda*, chapter VII: Human-Centric Adaptation, pages 71–72. Institute for Pervasive Computing, Johannes Kepler University Linz, Austria.

Omicini, A. and Denti, E. (2001).
From tuple spaces to tuple centres.
Science of Computer Programming, 41(3):277–294.

Omicini, A. and Viroli, M. (2011).
Coordination models and languages: From parallel computing to self-organisation.
The Knowledge Engineering Review, 26(1):53–59.
Special Issue 01 (25th Anniversary Issue).

References V

Omicini, A. and Zambonelli, F. (1999).
Coordination for Internet application development.
Autonomous Agents and Multi-Agent Systems, 2(3):251–269.
Special Issue: Coordination Mechanisms for Web Agents.

Viroli, M. and Casadei, M. (2009).
Biochemical tuple spaces for self-organising coordination.
In Field, J. and Vasconcelos, V. T., editors, *Coordination Languages and Models*, volume 5521 of *LNCS*, pages 143–162. Springer, Lisbon, Portugal.
11th International Conference (COORDINATION 2009), Lisbon, Portugal, June 2009. Proceedings.

Weyns, D., Omicini, A., and Odell, J. J. (2007).
Environment as a first-class abstraction in multi-agent systems.
Autonomous Agents and Multi-Agent Systems, 14(1):5–30.
Special Issue on Environments for Multi-agent Systems.

References VI

Zambonelli, F., Castelli, G., Ferrari, L., Mamei, M., Rosi, A., Di Marzo, G., Risoldi, M., Tchao, A.-E., Dobson, S., Stevenson, G., Ye, Y., Nardini, E., Omicini, A., Montagna, S., Viroli, M., Ferscha, A., Maschek, S., and Wally, B. (2011).
Self-aware pervasive service ecosystems.
Procedia Computer Science, 7:197–199.
Proceedings of the 2nd European Future Technologies Conference and Exhibition 2011 (FET 11).

$\mathcal{M}$olecules *of* $\mathcal{K}$nowledge

Self-Organisation in Knowledge-Intensive Environments

Stefano Mariani

s.mariani@unibo.it

DISI

Alma Mater Studiorum—Università di Bologna

Seminar @ Distributed Systems Group

TU Wien, Austria - 29th April 2014