Self-Organising News Management: The Molecules of Knowledge Approach

Stefano Mariani, Andrea Omicini {s.mariani, andrea.omicini}@unibo.it

ALMA MATER STUDIORUM—Università di Bologna

ASENSIS 2012 ASENSIS Workshop @ SASO Conference Lyon, 10th of September 2012



1 / 33

Mariani, Omicini (Università di Bologna)

Self-Organising News Management

Motivations

The Molecules of Knowledge model
 Formal MOK

• MoK in SAPERE

The news management scenario
 IPTC standards

- IPIC standard
- MoK-News





Outline

Motivations

- 2 The Molecules Of Knowledge model
 - Formal MoK
 - MoK in SAPERE
- 3 The news management scenario
 - IPTC standards
 - MoK-News





イロト イヨト イヨト イヨト

The challenge

Knowledge-intensive environments...

...present new critical challenges in the *knowledge management process*: the ever-increasing amount of information to handle, its heterogeneity in structure, and the pace at which it is made available are just a few to mention.

For journalists, in particular...

...today ICT systems provide at the same time new opportunities and new obstacles: the ability to find all the *relevant information* needed in the short time being a issue that even the most advanced general-purpose research engines are not able to face.



A Tuple-based answer

Adaptive and self-organising systems...

...seem the only possible answer when the scale of the problem is too huge, unpredictability too high, global control unrealistic, and deterministic solutions simply do not work [1].

Biochemical tuple spaces...

...push tuple-based coordination models toward self-organisation by exploiting the *chemical metaphor* enhanced with topology aspects [2]:

- $\rightarrow\,$ tuples are seen as chemical reactants possibly diffusing through chemical membranes resembling a network of cells
- → tuple spaces act as *chemical solutions simulators*, that is evolve tuples concentrations ^a exactly as chemical reactants do in chemical substances following the well-known Gillespie algorithm [3].

^atheir relative quantity w.r.t. the others

Goals

On one hand...

...to bring the biochemical tuple space abstraction and its self-organising features to its full realization into *knowledge intensive environments* – in particular, news management systems – so to harness their complexity.

On the other hand...

...to keep some form of "backward compatibility" with knowledge representation & sharing standards so to smooth the gap from legacy models and systems — again, focussing on news industry.



6 / 33

イロト 不得下 イヨト イヨト

Outline

Motivations

2 The Molecules of Knowledge model

- Formal MoK
- MoK in SAPERE
- 3 The news management scenario
 - IPTC standards
 - MoK-News





Aims & inspiration

The MOK model...

...features the idea that knowledge should autonomously *aggregate* and *diffuse* to reach knowledge consumers rather than be "searched" [4].

Biochemical tuple spaces...

...seem the most promising abstraction to bring such self-organising features into distributed knowledge intensive systems.



MoK abstractions I

MoK main abstractions are

atoms the smallest *unit of knowledge* in MoK, contain information from a source and belong to a compartment — thus being subject to its "laws of nature"

molecules the ${\rm MOK}$ units for knowledge aggregation, bond together "somehow-related" atoms

enzymes emitted by MOK catalysts, represent *prosumer's actions* and participate MOK reactions to *affect* the way in which atoms and molecules evolve

reactions working at a given *rate* ^{*a*}, they regulate the evolution of each MOK compartment, by ruling the way in which molecules *aggregate*, *be reinforced*, *diffuse*, and *decay*

^aaffected by molecules concentrations

${\rm MoK}$ abstractions II

MoK other abstractions are

compartments the spatial abstraction of MOK, compartments represent the conceptual loci for all MOK entities as well as for MOK biochemical processes – that is, reaction execution –, and provide MOK with the notions of *locality* and *neighbourhood*

sources each one associated to a compartment, MoK sources are the *origins of knowledge*, which is continuously injected at a certain *rate* in the form of MoK atoms

catalysts the abstraction for *knowledge prosumers*, catalysts emit enzymes in order to attract to him/her relevant knowledge items



Envisioning MOK systems

A MoK system...

...should be seen as a *network of biochemical compartments* – distributed shared information spaces – in which sources of knowledge continuously *inject atoms of knowledge*; these may then *aggregate in molecules* – more complex knowledge chunks – and *diffuse* toward other biochemical compartments — depending on biochemical reactions.

Knowledge producers, managers and consumers...

...are modelled as catalysts, whose information-oriented actions are *reified in terms of enzymes* influencing atoms aggregation within their own workspaces – mapped into biochemical compartments – and molecules diffusion toward other catalysts compartments.



A MOK system



Formal MOK

MoK formal model

atom(src, val, attr) | molecule(Atoms) | enzyme(Atoms)

Reactions semantics

Abstractions syntax

 $molecule(Atoms_1) + molecule(Atoms_2) \mapsto^{r_agg}$ $molecule(Atoms_1 | Atoms_2) + Residual(Atoms_1, Atoms_2)$

$$\begin{array}{rl} \texttt{enzyme}(\texttt{Atoms}_1) \ + \ \texttt{molecule}(\texttt{Atoms}_2)_c \longmapsto^{\texttt{r_reinf}} \\ \texttt{molecule}(\texttt{Atoms}_2)_{c+1} \end{array}$$

 $molecule(Atoms)_{c} \mapsto^{r_{decay}} molecule(Atoms)_{c-1}$

 \parallel Molecules₁ \bigcup molecule₁ $\parallel_{\sigma^i} + \parallel$ Molecules₂ $\parallel_{\sigma^{ii}} \rightarrow^{r_diffusion}$ $\|$ Molecules₁ $\|_{\sigma^i} + \|$ Molecules₂ [molecule₁ $\|_{\sigma^{ii}}$

MOK and SAPERE coexistence I

SAPERE in short

SAPERE is a model for the engineering of complex self-organising and adaptive pervasive service ecosystems [5]. Inspired by biochemical tuple spaces, in SAPERE agents share LSAs (Live Semantic Annotation) - kind of tuples – representing them in shared contexts and allowing them to interact and pursue their own goals. LSAs are managed by SAPERE eco-laws - kind of chemical-like rules - which are responsible to evolve LSAs according to both agents' and system's needs.

The \overline{MOK} model...

...could be placed within the SAPERE model because

! MoK doesn't make any *technical assumption* about sources, compartments and catalysts neither

MOK and SAPERE coexistence II

Hence...

 \dots the MOK model could be used as a part of the SAPERE model: the former *focusses on information* whereas the latter on everything else typically, agents & services

A mapping...

... between models could then be devised:

- \mapsto MoK atoms and enzymes could be implemented as SAPERE LSAs
- \mapsto MOK reactions as SAPERE eco-laws
- \mapsto MoK molecules could be reconstructed by navigating SAPERE bonds between LSAs — or be reified as LSAs as well
- \mapsto MoK compartments could be mapped to SAPERE nodes
- \mapsto MOK sources & catalysts to SAPERE agents

A B A A B A

Image: A matrix

Outline

Motivations

- 2 The Molecules Of Knowledge model
 - Formal MoK
 - MoK in SAPERE

The news management scenario

- IPTC standards
- MoK-News

4 Conclusions & future works



16 / 33

イロト イヨト イヨト イヨト

Why news

News management systems...

... are a prominent example of:

heterogeneity News sources can be virtually anything, from handwritten notes to printed official documents through web published articles

ubiquity Netbooks, tablets and smartphones pushed information production, sharing and consumption to be *pervasive* as never before

unpredictability News producers are no longer graduated journalists solely, they include bloggers and whoever has access to the web though



NewsML & NITF |

The IPTC...

...is a consortium of the world's major news agencies, news publishers and news industry vendors ^a. IPTC *develops and maintains technical standards for improved news management* — used among the other by the italian ANSA, the american Thomson Reuters, and the english BBC.

^ahttp://www.iptc.org/site/Home/About/

NewsML

The NewsML tagging language ^a is a media-type orthogonal news sharing format standard aimed at conveying not only the core news content, but also the data that describe the content in an abstract way — that is *metadata*.

^ahttp://www.iptc.org/site/News_Exchange_Formats/NewsML-G2/

NewsML & NITF ||

XML & NewsCodes

In order to ease syntactical and semantical interoperability, NewsML adopts XML as the first implementation language for its standards and maintains sets of *controlled vocabularies* – collectively branded as NewsCodes ^a – to represent concepts describing and categorising news objects in a consistent manner — pretty much as domain-specific *ontologies* do.

^ahttp://iptc.cms.apa.at/site/NewsCodes/



NewsML & NITF III

NITF

The News Industry Text Format ^a enable journalists to *enrich the content* of news articles, supporting the identification and description of a number of news typical features, among which the most notable are: Who owns the copyright to the item, who may republish it, and who it's about What subjects, organisations, and events it covers When it happened, was reported, issued, and revised Where it was written, where the action took place, and where it may be released

The NITF too adopts both XML and NewsCodes.

^ahttp://www.iptc.org/site/News_Exchange_Formats/NITF/

MOK-News

Toward MOK-News I

A mapping...

 \dots from news representation standards to the MOK model can be devised to specialize the MoK model upon the news management scenario.

MoK atoms in fact...

...have a clear counterpart in NewsML & NITF: the tag. Tags - along with their "content" – can be seen as the atoms that altogether compose the "news-substance" — that is an article, a blog post, or whatever the like.



21 / 33

Toward $\operatorname{MoK-News}$ []

Therefore...

 $\dots {\rm MoK}$ molecules and biochemical reactions – and enzymes, too – are both syntactically and semantically affected by the new ${\rm MoK-News}$ atoms. Here, in fact, molecules can be re-interpreted as *ever-growing news pieces*, and biochemical reactions as *news aggregators* and *diffusers*.

E.g.

A MOK reaction could exploit the catalog information – pointing to a NewsCodes ontology – to check if atoms content is semantically related, then, aggregate them.



A B b

$\operatorname{MoK-News}$ formal model

Formally...

```
...a generic {\rm MOK} atom of the form atom(src, val, attr)_c becomes a specialised {\rm MOK}\textsc{-News} atom of the form
```

```
atom(src, val, sem(tag, catalog))<sub>c</sub>
```

where

```
src ::= news source uri
val ::= news content
attr ::= sem(tag, catalog)
   tag ::= NewsML tag | NITF tag
   catalog ::= NewsCode uri | ontology uri
```



23 / 33

A B F A B F

MoK-News

Envisioning MOK-News systems I

A MoK-News systems...

...should hence be seen as a *self-organising news repository* in which:

- ! news pieces "tag-content" pairs are injected either automatically - e.g. using XML parsers - or manually - by journalists - in the form of MoK-News atoms
- ! enzymes are released by catalysts journalists as manifestations of their actions over knowledge
- ! biochemical reactions

aggregate together *semantically related* atoms — based upon catalog information

diffuse atoms/molecules in neighborhood compartments reinforce them by using enzymes

decay non-relevant information

24 / 33

3

Envisioning MoK-News systems II

"Smart diffusion" ...

... is achieved as a *self-organising* process caused by the cooperation among diffusion, reinforcement – of relevant knowledge, that is more frequently accessed – and decay — of useless information, ignored by catalysts.

E.g.

A journalist interested in sports news is *more likely* to search, read, annotate – generally, access – sport-related atoms. In the process, he/she releases enzymes which reinforce accessed atoms/molecules concentration. In the very end, his/her compartment will mainly store sports-related knowledge.



25 / 33

Envisioning $\operatorname{MOK-News}$ systems III

00	Engine Console Agent	
Tite generator		
Operation	•	
out(law([atom(SRC,VAL,ATTR)], 0.1, [iring(atom(SRC,VAL,ATTR))]))	
● ● ●	Engine Console Agent	
Titleconomics		
Operation		÷
<pre>out(low([srzyme(atom(" ipric)</pre>	<pre>Jobl:", val('Treasury News', "AMCG: TRS'), val('Treasury News', "AMCG: TRS'), 'val('Treasury News', Trans'), 'val('treasury News', trans'), 'val('t</pre>	
000	Engine Console Agent	
e sports		
Operation	and Maharahi Indonesianah and S	
<pre>reactant(atom()ptcID.0111, vu()'The 2 sem(creator, 'http://www.xm reactant(atom()ptcID.0012, vu()'Vare's reactant(atom()ptcID.0011, vu()'Treas reactant(atom()ptcID.0112, vu()'<u>Graus</u> sem(genre, 'http://www.xml reactant(atom()ptcID.0112, vu()'<u>Arizs</u></pre>	<pre>parts_Network, 'web:sportsnetwork.com'), parts_Network, 'web:sportsnetwork.com'), ty Of Recessionary Forces Leave Las Vegas Economy Scarred', '),sem(headline, '')) (1) preview', 'spfixt:pre-event-coverage'), am.com/specification/xts-SportSddesCatalog_1.xml'))(37) na Diamondbacks', 'teami1.mlb.com-t.26'), team.com/specification/xts-SportSddesCatalog_1.xml'))(37)</pre>	
Omicini (Università di Bologn	Self-Organising News Management ASENSIS 2012, 10/9/2012	

Outline

Motivations

- 2 The Molecules Of Knowledge model
 - Formal MoK
 - MoK in SAPERE
- 3 The news management scenario
 - IPTC standards
 - MoK-News

4 Conclusions & future works



Final remarks

$Molecules \ Of \ Knowledge$

The MoK model

- $\rightarrow\,$ provides knowledge workers in general with a novel approach both in thinking and managing knowledge
- \rightarrow supports their work with *self-organising knowledge workspaces* [6] to interact with

$\operatorname{MoK-News}$

 ${\rm MoK}\textsc{-News}$ is a domain-specific instantiation of ${\rm MoK}$ focussed on news management systems, which

- \rightarrow assists professional journalists in their search for relationships between different information chunks potentially leading to a news story through aggregation
- ightarrow promotes news sharing toward interested readers through *diffusion*

Open issues & further developments

How to...

- ? ...push the MOK model toward the idea of *self-organising workspace* [6], fully supporting adaptiveness of compartments rather than information solely?
- ? ...to effectively implement such a model upon a real infrastructure dealing with real-world scenarios?

Further works...

- ...will be devoted to
 - ! explore techniques and methodologies to address first issue
 - improve current implementation of the MoK model upon TuCSoN
 [7] coordination infrastructure.



Thanks

Thanks to...

- ...everybody here for listening
- ...the SAPERE team for bringing me here ¹

¹This work has been supported by the EU-FP7-FET Proactive project SAPERE Self-aware Pervasive Service Ecosystems, under contract no.256873.

Mariani, Omicini (Università di Bologna)

Self-Organising News Management

Bibliography I

A. Omicini and M. Viroli.

"Coordination models and languages: From parallel computing to self-organisation". The Knowledge Engineering Review, vol. 26, no. 1, pp. 53–59, Mar. 2011.



M. Viroli and M. Casadei.

"Biochemical tuple spaces for self-organising coordination".

Coordination Languages and Models, ser. LNCS, J. Field and V. T. Vasconcelos, Eds. Lisbon, Portugal: Springer, Jun. 2009, vol. 5521, pp. 143–162.



Gillespie, D.T.

"Exact stochastic simulation of coupled chemical reactions". The Journal of Physical Chemistry **81**(25), 2340–2361 (1977).

S. Mariani and A. Omicini.

"Molecules of Knowledge: Self-organisation in knowledge-intensive environments". Intelligent Distributed Computing VI, ser. Studies in Computational Intelligence. Calabria, Italy: Springer, 24-26 Sep. 2012.



Bibliography

Bibliography II

F. Zambonelli, G. Castelli, L. Ferrari, M. Mamei, A. Rosi, G. Di Marzo, M. Risoldi, A.-E. Tchao, S. Dobson, G. Stevenson, Y. Ye, E. Nardini, A. Omicini, S. Montagna, M. Viroli, A. Ferscha, S. Maschek, and B. Wally.

"Self-aware pervasive service ecosystems".

Procedia Computer Science, vol. 7, pp. 197–199, Dec. 2011, proceedings of the 2nd European Future Technologies Conference and Exhibition 2011 (FET 11).

A. Omicini.

"Self-organising Knowledge-intensive Workspaces".

Pervasive Adaptation. The Next Generation Pervasive Computing Research Agenda, Ch. 7, pages 71-72, May 2011.

Alois Ferscha (eds.), Institute for Pervasive Computing, Johannes Kepler University Linz, Austria.

A. Omicini, F. Zambonelli.

"Coordination for Internet Application Development".

Autonomous Agents and Multi-Agent Systems, Special Issue: Coordination Mechanisms for Web Agents (1999)



32 / 33

(日) (同) (三) (三)

Self-Organising News Management: The Molecules of Knowledge Approach

Stefano Mariani, Andrea Omicini {s.mariani, andrea.omicini}@unibo.it

ALMA MATER STUDIORUM—Università di Bologna

ASENSIS 2012 ASENSIS Workshop @ SASO Conference Lyon, 10th of September 2012



33 / 33

Mariani, Omicini (Università di Bologna)

Self-Organising News Management