Applications of P systems in population biology and ecology

P. Milazzo

joint work with
R. Barbuti, P. Bove and G. Pardini

1. Dept. of Computer Science, Univ. of Pisa, Italy.
2. Museum of Natural History, Univ. of Pisa, Italy.
Population modelling: motivations

- Models contribute to understanding the factors governing population growth, evolution, extinction, ...
  - Hypothesis validation
- Models allow making predictions on the future of a population of interest
  - e.g. endangered species
- Models can support decision making in planning control policies
  - e.g. reintroduction actions
Population modelling: application domains

- **Population biology**
  - causes of extinction of species, demography, ... 
- **Ecology**
  - sustainable development, biodiversity, ... 
- **Evolutionary biology**
  - species formation, ... 
- **Social sciences**
  - social behaviours, animal sociology, ... 
- **Epidemiology**
  - spread of diseases, role of vaccination, ...
Population modelling: traditional methodologies

• Mathematical modelling (ODEs, recurrence eq., …)
  – e.g. Lotka-Volterra predator-prey equations
  – e.g. Susceptible/Infected/Recovered (SIR) epidemic model
  – Problems: unfriendly notation, deterministic dynamics

• Agent based modelling
  – Individuals models as agents whose behaviour is described by an algorithm or set of rules
  – Probably the most used methodology in ecological modelling (Individual Based Modelling – IBM)
  – Problems: often unformalized/ambiguous
Population modelling: P systems

- P systems can provide a simple, elegant and unambiguous notation for population modelling.

- **Objects** can represent:
  - individuals (and their current state)
  - available natural resources (e.g. food)
  - state of the environment (e.g. season, weather)

- **Evolution rules** can represent events like:
  - birth, mating, oviposition, growth, death, predation, transmission of diseases, fight, communication, aggression, ...
Population modelling: P systems

• **Maximal parallelism** is good for modelling populations that **evolve by stages**
  – All the individuals are involved in the same activity (e.g. reproduction season, hibernation, ...)

• Particularly useful if combined with **rule promoters**
  – to enable different rules during different stages

• But also **probabilities** are necessary
  – sometimes individuals can be subject to alternative events (e.g. birth of male/female), or can make choices
  – in particular when the population size can be small
Minimal Probabilistic P systems

- These observations led us to the definition of Minimal Probabilistic P systems (MPP systems).
- They are P systems based on *Probabilistic maximal parallelism with rule promoters*.
- They are minimal in the sense that we tried to include as less features as possible...
- No membrane structure...
Minimal Probabilistic P systems

**MPP system** A *Minimal Probabilistic P system* is a tuple \( \langle V, w_0, R \rangle \) where:

- \( V \) is a possibly infinite alphabet of objects, with \( V^* \) denoting the universe of all multisets having \( V \) as support.
- \( w_0 \in V^* \) is a multiset describing the initial state of the system.
- \( R \) is a finite set of evolution rules having the form

\[
  u \xrightarrow{f} v |_p
\]

where \( u, v, p \in V^* \) are multisets (often denoted without brackets) of reactants, products and promoters, respectively, and \( f : V^* \mapsto \mathbb{R}^{\geq 0} \) is a rate function.
Probabilistic maximal parallelism

Briefly: pick rules one-by-one with probabilities proportional to their rates until you get a maximal multiset of rule instances

---

**Algorithm 1 Probabilistic maximally parallel evolution step**

```plaintext
function STEP(\(w\))
    \(x = w\)
    \(y = \emptyset\)

    while there exists \(u \xrightarrow{f} v \mid p \in R\) s.t. \(u \subseteq x\) and \(p \subseteq w\) do
        \(R' = \{u \xrightarrow{f} v \mid p \in R \mid u \subseteq x\) and \(p \subseteq w\}\)
        choose \(u' \xrightarrow{f'} v' \mid p' \) from \(R'\) with a probability proportional to \(f'(w)\)
        \(x = x \setminus u'\)
        \(y = y \cup v'\)
    end while

    return \(x \cup y\)

end function
```
Probabilistic maximal parallelism

• In the end, probabilistic maximal parallelism turns out to use probabilities just to choose among rules that compete for the same objects.

• An applicable rule that does not compete with any other rule will be for sure applied, whatever its rate is.

• **Note:** applicable rules should always have a positive rate.
Analysis techniques

• Simulation

• Statistical model checking:
  – The analysis technique we choose (and suggest) for population and ecosystem modelling
  – A statistical model checker:
    1. runs a number of simulations of the model of interest
    2. use simulation results (execution traces) to construct a Discrete Time Markov Chain representing the system behaviours
    3. verifies behavioural properties (expressed as temporal logic formulas) on the Markov Chain (model checking)

– We defined the translation of MPP systems into the PRISM (model checker) input language
Application: hybrid populations of water frogs

• We applied MPP systems to investigate an open problem in evolutionary biology:
  – To understand the mechanisms underlying the stability of European hybrid populations of water frogs
Among European water frogs there are two species ...

*Pelophylax lessonae*

Adapted to mashes and ponds
Pool frogs

*Differences*

Vocal sacs

*Pelophylax ridibundus*

Adapted to lakes
Lake frogs

*Size*
which interbred producing hybrids with intermediate characteristics

**Pelophylax lessonae**

**Pelophylax ridibundus**

**Pelophylax esculentus**

*Note: for size reasons the interbreeding involves* **P. lessonae** *males and P. ridibundus* *females*
Some notation

Male sex chromosome

Pelophilax lessonae

Pelophilax ridibundus

Pelophilax esculentus

Males and females

L_yL

L_yR  LR

RR
*P. ridibundus* are currently limited to Eastern Europe.

In Western Europe are diffused populations of coexisting *P. lessonae* and *P. Esculentus*

**L-E complexes**
*P. esculentus* have a particular gametogenesis (hemiclonal)

*Pelophilax lessonae*

*Pelophilax esculentus*

Hemiclonality: there is no recombination between chromosomes
Resulting in the following reproduction table

<table>
<thead>
<tr>
<th></th>
<th>LL</th>
<th>LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>L_yL</td>
<td>L_yL</td>
<td>L_yR</td>
</tr>
<tr>
<td>L_yR</td>
<td>LR</td>
<td>RR</td>
</tr>
</tbody>
</table>
Resulting in the following reproduction table

<table>
<thead>
<tr>
<th></th>
<th>LL</th>
<th>LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_yL$</td>
<td>$L_yL$</td>
<td>$L_yR$</td>
</tr>
<tr>
<td>$L_yR$</td>
<td>LL</td>
<td>LR</td>
</tr>
</tbody>
</table>

usually inviable
Resulting in the following reproduction table

<table>
<thead>
<tr>
<th></th>
<th>L_yL</th>
<th>LL</th>
<th>L_yR</th>
<th>LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>L_L</td>
<td><img src="image1.png" alt="Frog" /></td>
<td><img src="image2.png" alt="Frog" /></td>
<td><img src="image3.png" alt="Frog" /></td>
<td><img src="image4.png" alt="Frog" /></td>
</tr>
<tr>
<td>L_R</td>
<td><img src="image5.png" alt="Frog" /></td>
<td><img src="image6.png" alt="Frog" /></td>
<td><img src="image7.png" alt="Frog" /></td>
<td><img src="image8.png" alt="Frog" /></td>
</tr>
</tbody>
</table>

Numerical advantage for hybrids: L_yL

Usually inviable: RR
Consequences:
- Hybrids are numerically advantaged
- Hybrids show heterosis (hybrid vigor)

They should outcompete the parent species (*P. lessonae*), but *P. esculentus* alone cannot survive!! (they can survive only as reproductive parasites)
How can L-E complexes not to get extinct?
An answer based on observations and experiments:

**female sexual preferences**

In water frogs females are choosy and males are promiscuous.

**lyL**

**LL**

**LyR**

**LR**

females of both species prefer *P. lessonae* males
Consequence of female preferences:

If female preferences are strong enough, this entry is negligible.
Consequence of female preferences:

If female preferences are strong enough this entry is negligible.
Why *P. ridibundus* are not viable?

- The R genome is transmitted clonally, without any recombination, thus:
  - accumulated deleterious mutations cannot be purged
  - this phenomenon is known as “Muller’s ratchet”.

In hybrids the dysfunctionalities of the R genome are compensated by the L genome
L-E complexes: MPP systems model

The model of L-E complexes is the MPP systems \( \langle V_{LE}, w_{0LE}, R_{LE} \rangle \)

where \( V_{LE} = V_{LEa} \cup V_{LEj} \cup V_{ctrl} \)

with

\[
V_{LEa} = \{ LL, L_yL, LR_*, L_yR_*, LR_0, L_yR_0, R_*R_0, R_0R_0 \}
\]

\[
V_{LEj} = \{ LL^j, L_yL^j, LR_*^j, L_yR_*^j, LR_0^j, L_yR_0^j, R_*R_*^j, R_*R_0^j, R_0R_0^j \}
\]

\[
V_{ctrl} = \mathbb{N} \cup \{ REPR, SEL \}
\]
L-E complexes: MPP systems model

Evolution rules:

**REPRODUCTION**

For each kind of male \( x \), female \( y \) and juvenile \( z \):

\[
\begin{align*}
  x \ y \overset{f_{xy}}{\longrightarrow} x \ y \ z \mid_{REPR}
\end{align*}
\]

where:

\[
f_{xy}(w) = k_{mate}(x, y) \cdot |w|_x \cdot |w|_y \cdot 1/k_{o\_kind}(x, y)
\]
L-E complexes: MPP systems model

Evolution rules:

**SELECTION (AND GROWTH)**

For each kind of individual $x$ and juvenile $x^j$:

\[
\begin{align*}
    x & \xrightarrow{g_x} x \mid SEL \\
    x^j & \xrightarrow{g_{xj}} x \mid SEL
\end{align*}
\]

\[
\begin{align*}
    x & \xrightarrow{g'_x} \epsilon \mid SEL \\
    x^j & \xrightarrow{g'_{xj}} \epsilon \mid SEL
\end{align*}
\]

where:

\[
g_x(w) = \frac{1}{\sigma + \frac{|w|}{k_{fit}(x) \cdot cc}}
\]

\[
g'_x(w) = 1 - g_x(w)
\]
L-E complexes: MPP systems model

Evolution rules:

STAGES ALTERNATION

\[ \text{REPR } 1 \rightarrow \text{REPR } 2 \]
\[ \text{REPR } 3 \rightarrow \text{SEL} \]
\[ \text{REPR } 2 \rightarrow \text{REPR } 3 \]
\[ \text{SEL} \rightarrow \text{REPR } 1 \]
In the end, the model description is rather compact...

\[
x \ y \xrightarrow{f_{xy}} \ x \ y \ z \mid \text{REPR}
\]

\[
x \xrightarrow{g_x} \ x \mid \text{SEL}
\]

\[
x^j \xrightarrow{g_{xj}} \ x \mid \text{SEL}
\]

\[
\text{REPR 1} \rightarrow \text{REPR 2}
\]

\[
\text{REPR 2} \rightarrow \text{REPR 3}
\]

\[
\text{REPR 3} \rightarrow \text{SEL}
\]

\[
\text{SEL} \rightarrow \text{REPR 1}
\]

\[
f_{xy}(w) = k_{mate}(x, y) \cdot |w|_x \cdot |w|_y \cdot 1/k_{o\_kind}(x, y)
\]

\[
g_x(w) = \frac{1}{\sigma + \frac{|w|}{k_{fit}(x) \cdot cc}}
\]

\[
g_x'(w) = 1 - g_x(w)
\]
Dynamics of a L-E complex (simulation)

- all R genomes have deleterious mutations
- the sexual preference for *P. lessonae* males is twice than that for *P. esculentus* males
- initial population: 95% of *P. lessonae* and 5% of *P. esculentus*
Probability of extinction

- Statistical model checking (1000 simulations)
- Probability of extinction in 60 years

\[ P = \text{? [} F \text{ total\_population=0 & years\_counter<=60} \text{]} \]

- Result: 0.01
What happens if *P. ridibundus* are viable?

The number of *P. esculentus* increases. *P. lessonae* decrease until their extinction. *P. esculentus* and *P. ridibundus* (females) cannot survive: they produce only *P. ridibundus* females.
Dynamics of a L-E complex (simulation)

- all R genomes are mutation-free
- the sexual preference for *P. lessonae* males is twice than that for *P. esculentus* males
- initial population: 95% of *P. lessonae* and 5% of *P. esculentus*
Dynamics of a L-E complex (simulation)

- all R genomes are mutation-free
- there is no sexual preference
- initial population: 95% of *P. lessonae* and 5% of *P. esculentus*
Statistical model checking: probabilities of *P. ridibundus* and Extinction

![Graph showing the probability of *P. ridibundus* and Extinction over years. The first graph peaks around the 20th year, while the second graph shows a steady increase towards 1.0 probability as years increase.]
In this scenario deleterious mutations are necessary for the stability of L-E complexes.

In all the existent Western Europe L-E complexes, generated *P. ridibundus* are inviable.
How L-E complexes react to the introduction of translocated *P. ridibundus*?
How L-E complexes react to the introduction of translocated *P. ridibundus*?

The result can be a monospecific *P. ridibundus* population...
How L-E complexes react to the introduction of translocated *P. ridibundus*?

But *P. ridibundus* can suffer for an unsuitable environment so that they can be quickly eliminated, but they can introduce mutations free genomes in the L-E complex.

... and eventually the whole population collapses.
Probability of extinction

- Statistical model checking (1000 simulations)

Result: 0.18
Conclusions

• P systems as an elegant notation for population models

• *Simulation and statistical model checking* as effective analysis techniques

• Case study on *lake frogs*: provided plausible answer to a currently open question in evolutionary biology

• **Further step**: Attributed Probabilistic P systems (APP systems) and their application to the modelling of social interactions in primates
References

R. Barbuti, P. Bove, P. Milazzo and G. Pardini
Minimal Probabilistic P systems for modelling ecological systems
Theoretical Computer Science 608 (2015) 36-56

P. Bove, P. Milazzo and R. Barbuti
The role of deleterious mutations in the stability of hybridogenetic water frog complexes
BMC Evolutionary Biology 14 (2014) 107

R. Barbuti, A. Bompadre, P. Bove, P. Milazzo and G. Pardini
Attributed Probabilistic P systems and their application to the modelling of social interactions in primates