Expressiveness Issues in Calculi for Artificial Biochemistry

$A \xrightarrow{r} C_1+\ldots+C_n$ $\Leftrightarrow$ $A ::= \tau@r;C_1|\ldots|C_n + b@s;\mathbf{0}$

$A+B \xrightarrow{s} D_1+\ldots+D_m$ $\Leftrightarrow$ $B ::= \overline{b}@s;D_1|\ldots|D_m$

What is the computational power of this calculus?

Gianluigi Zavattaro
University of Bologna

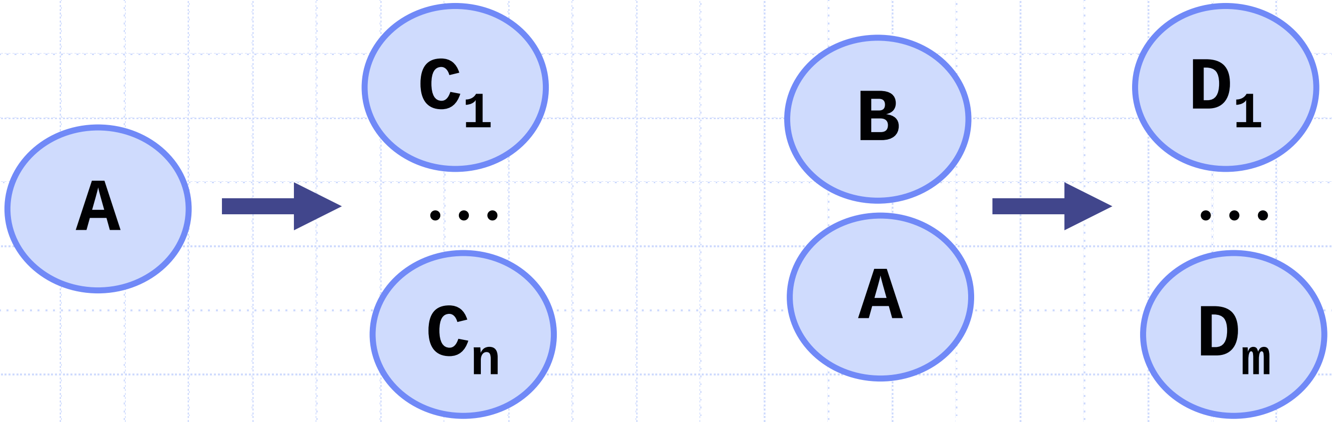
Based on joint work with
Luca Cardelli

Plan of the talk

- **Basic Chemistry and Basic Biochemistry**
  - **Biochemistry = Chemistry + complexation**
- Chemical Ground Form (CGF)
  - A process algebra for basic chemistry
- Biochemical Ground Form (BGF)
  - A process algebra for basic biochemistry
- Considered TERMINATION problems:
  - Existential termination in CGF (DECIDABLE)
  - Existential termination in BGF (UNDECIDIBLE)
  - Universal termination in CGF
    - Nondeterministic -all computations terminate- (DECIDABLE)
    - Probabilistic -terminate with probability 1- (UNDECIDABLE)

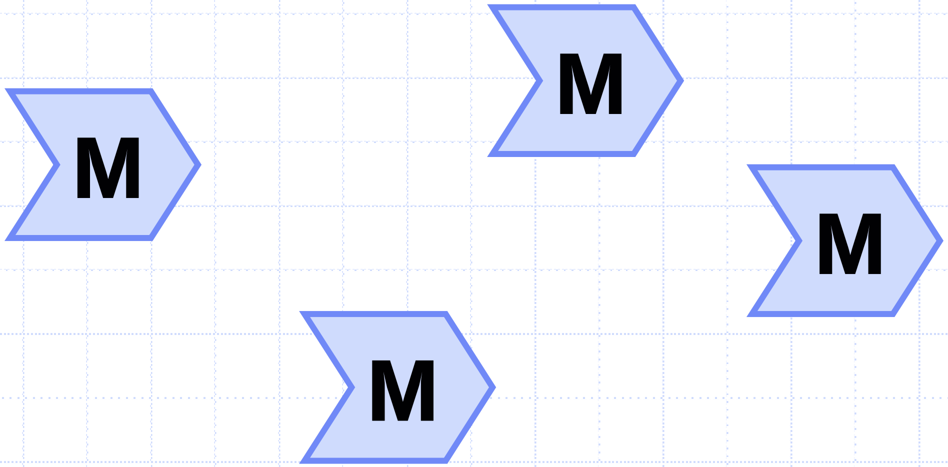
Basic Chemistry

- Molecules belong to Species
- Behavior described by reactions:
  - Monomolecular: $A \rightarrow C_1 + \ldots + C_n$
  - Bimolecular: $A + B \rightarrow D_1 + \ldots + D_m$



Basic Biochemistry

- Molecules form and modify complexes
  - by means of association and dissociation

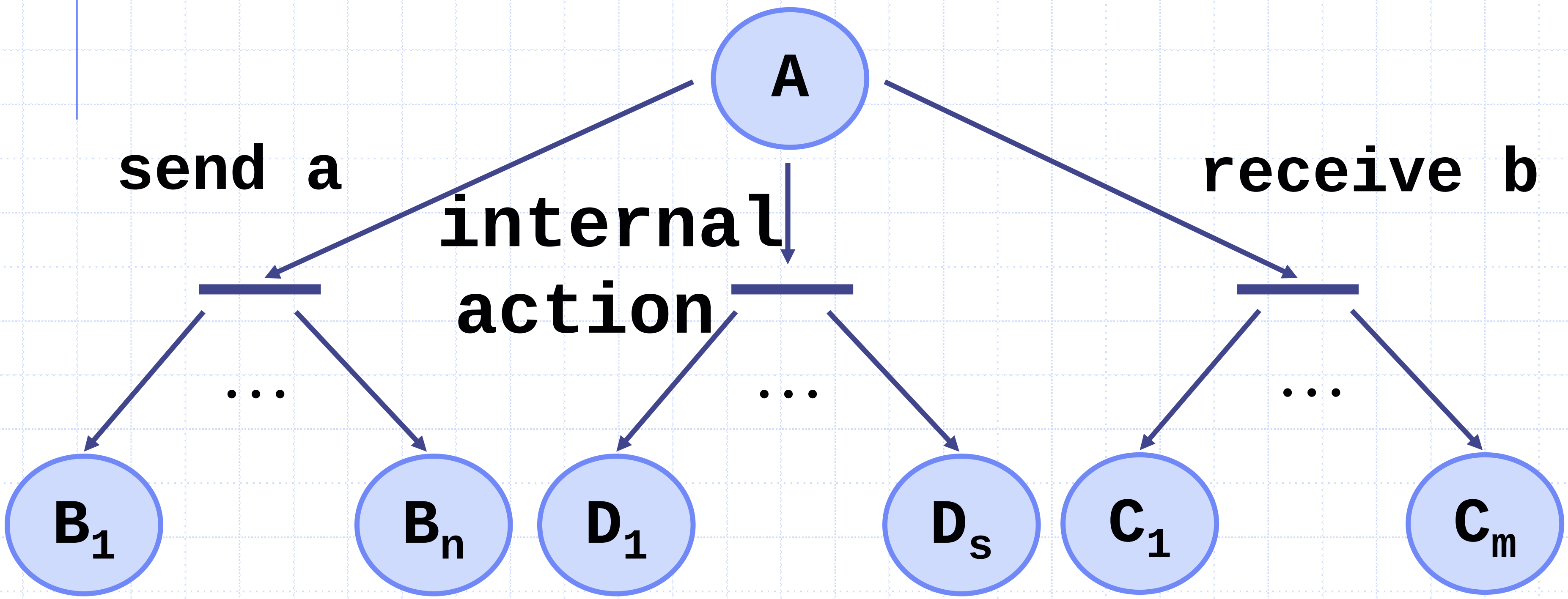


Plan of the talk

- Basic Chemistry and Basic Biochemistry
  - Biochemistry = Chemistry + complexation
- **Chemical Ground Form (CGF)**
  - **A process algebra for basic chemistry**
- Biochemical Ground Form (BGF)
  - A process algebra for basic biochemistry
- Considered TERMINATION problems:
  - Existential termination in CGF (DECIDABLE)
  - Existential termination in BGF (UNDECIDIBLE)
  - Universal termination in CGF
    - Nondeterministic -all computations terminate- (DECIDABLE)
    - Probabilistic -terminate with probability 1- (UNDECIDABLE)

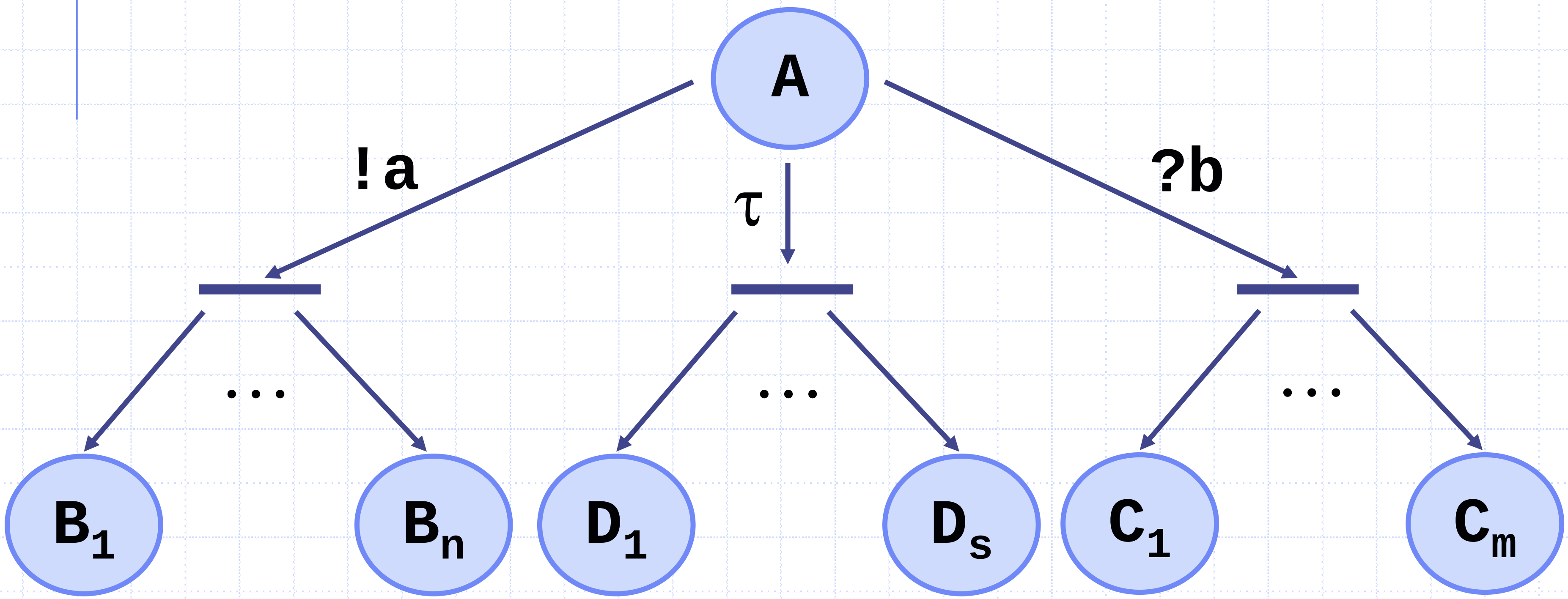
Chemical Ground Forms

- Stochastic variant of Milner's CCS, with an equivalent graphical notation (Stochastic Collective Automata)



Chemical Ground Forms

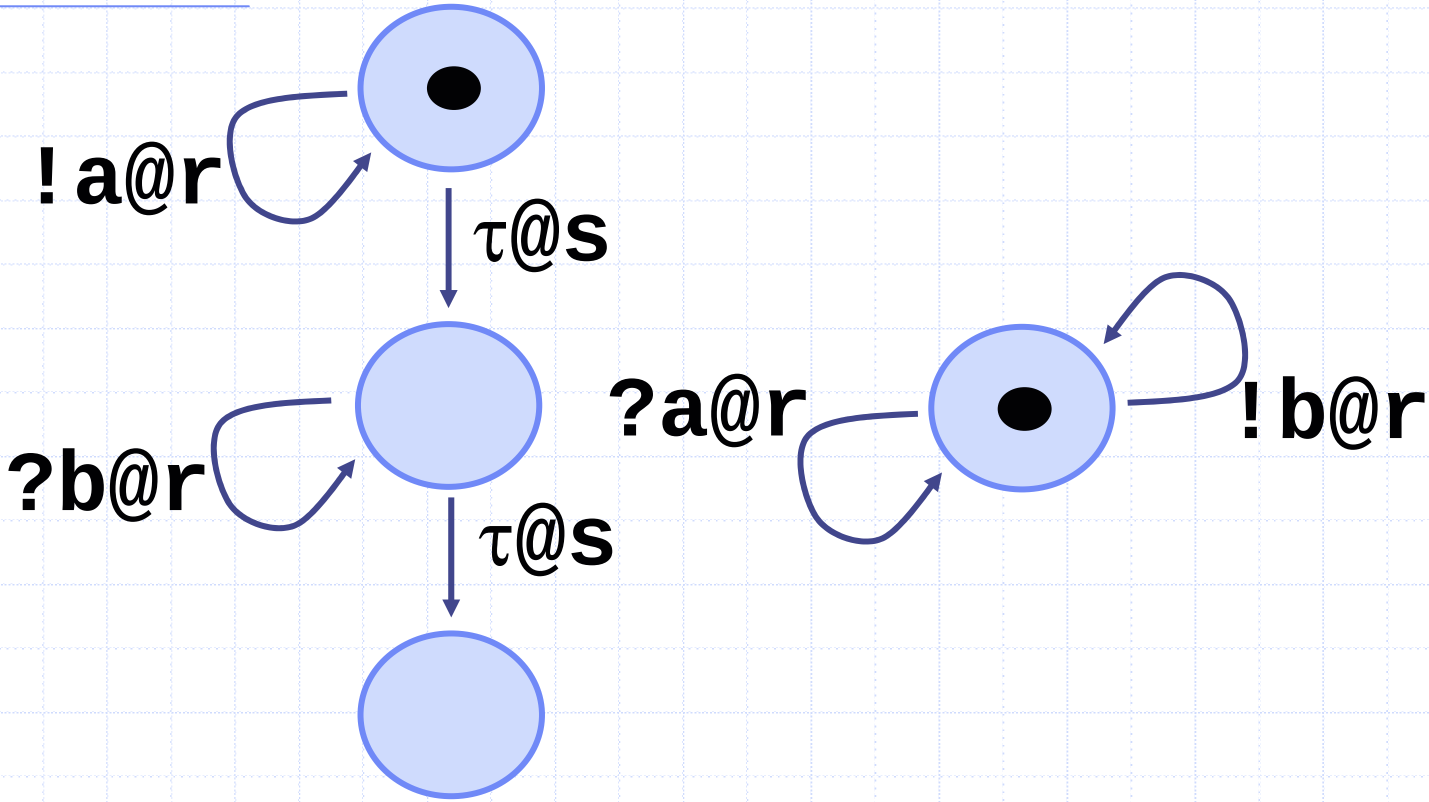
- Stochastic variant of Milner's CCS, with an equivalent graphical notation (Stochastic Collective Automata)



Why stochastic…

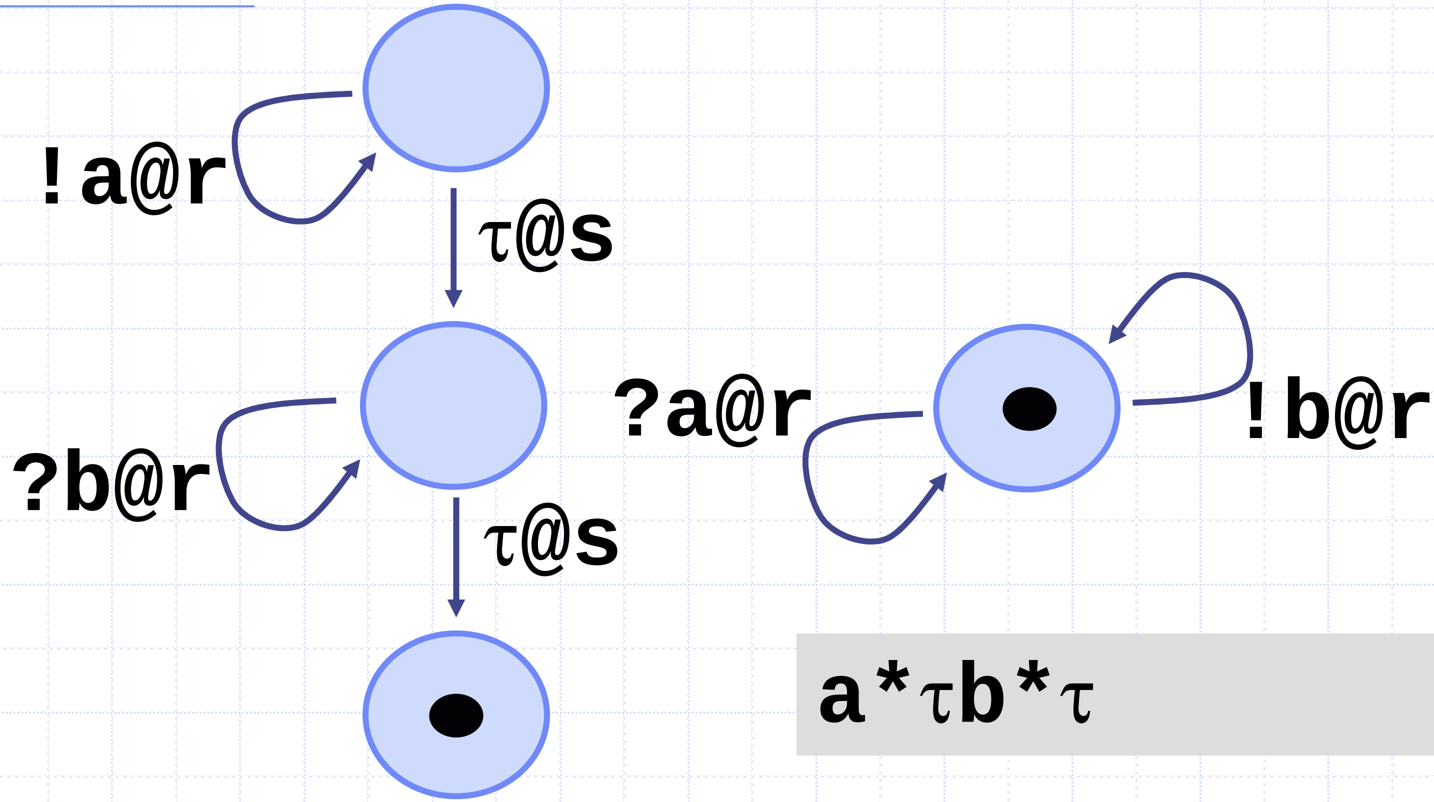
- Actions take (a variable amount of) time
- Each action has an associated rate **r**
  - Internal delay: τ@**r**
    - **Pr(internal delay < t) = 1-e**$^{-rt}$
  - Synchronization between complementary actions: **?a@r, !a@r**
    - **Pr(synchronization time < t) = 1-e**$^{-rt}$

Example 1



- Starting population: **A|A'**

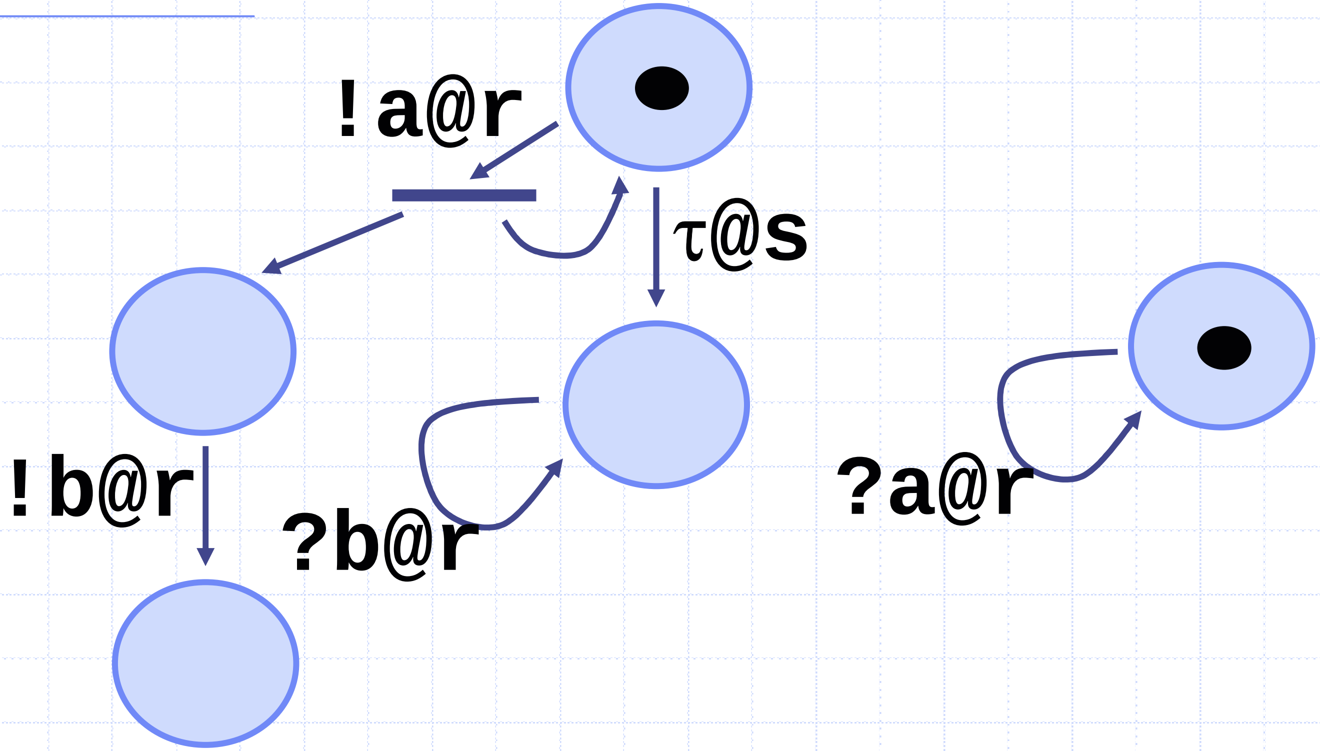
Example 1



a*τb*τ

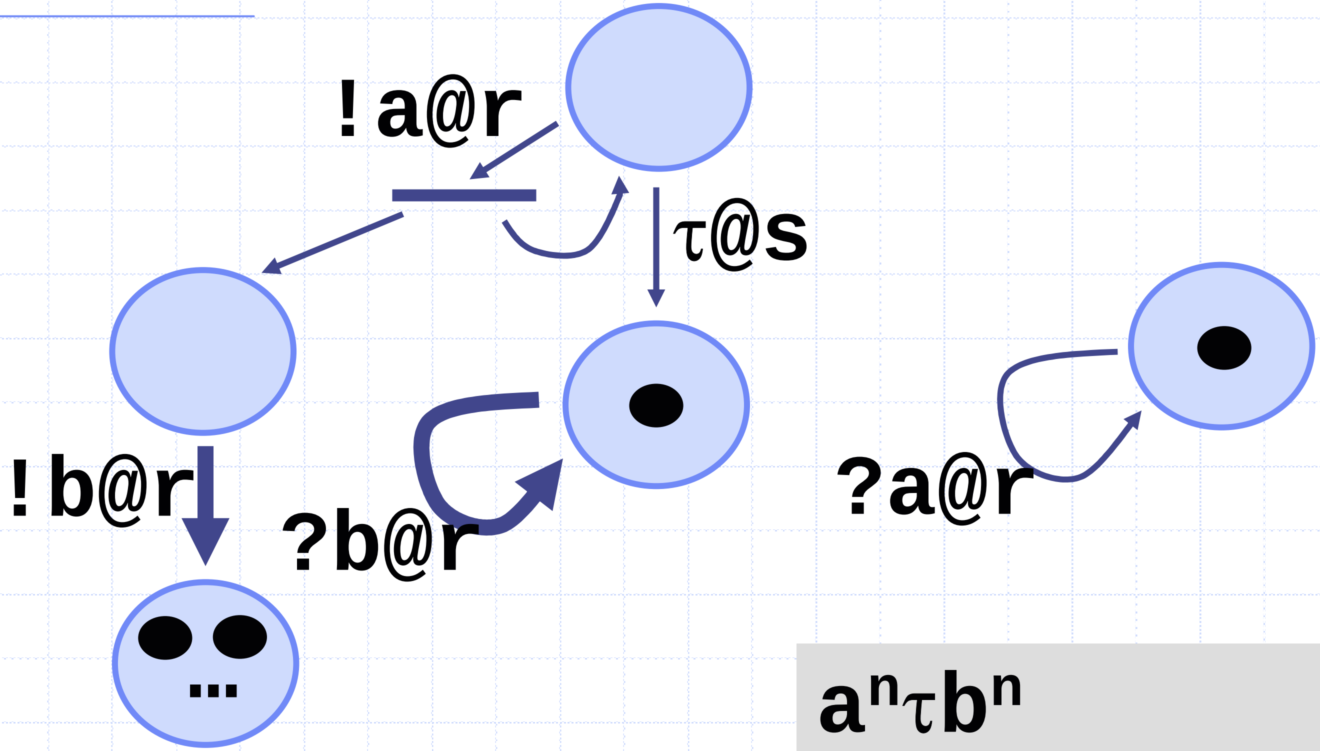
- Starting population: **A|A'**

Example 2



- Starting population: **A|A'**

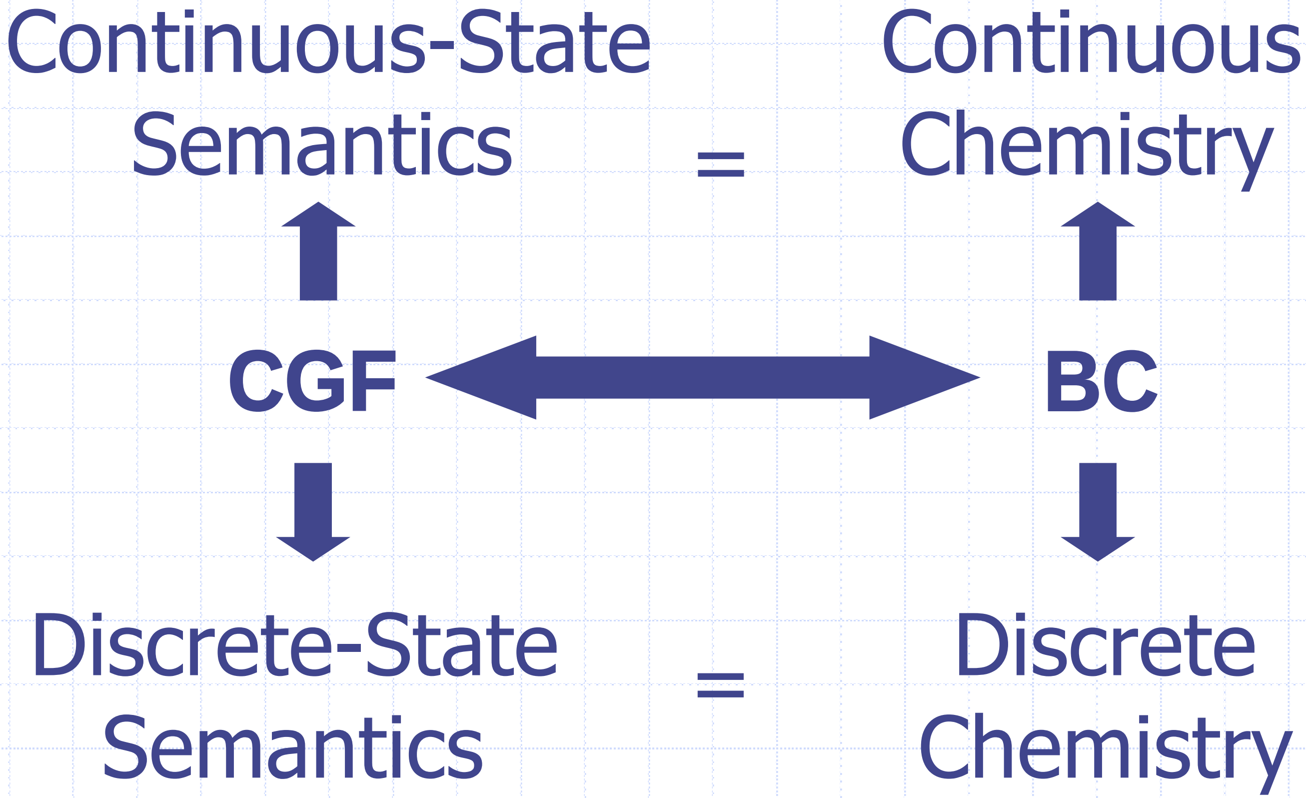
Example 2



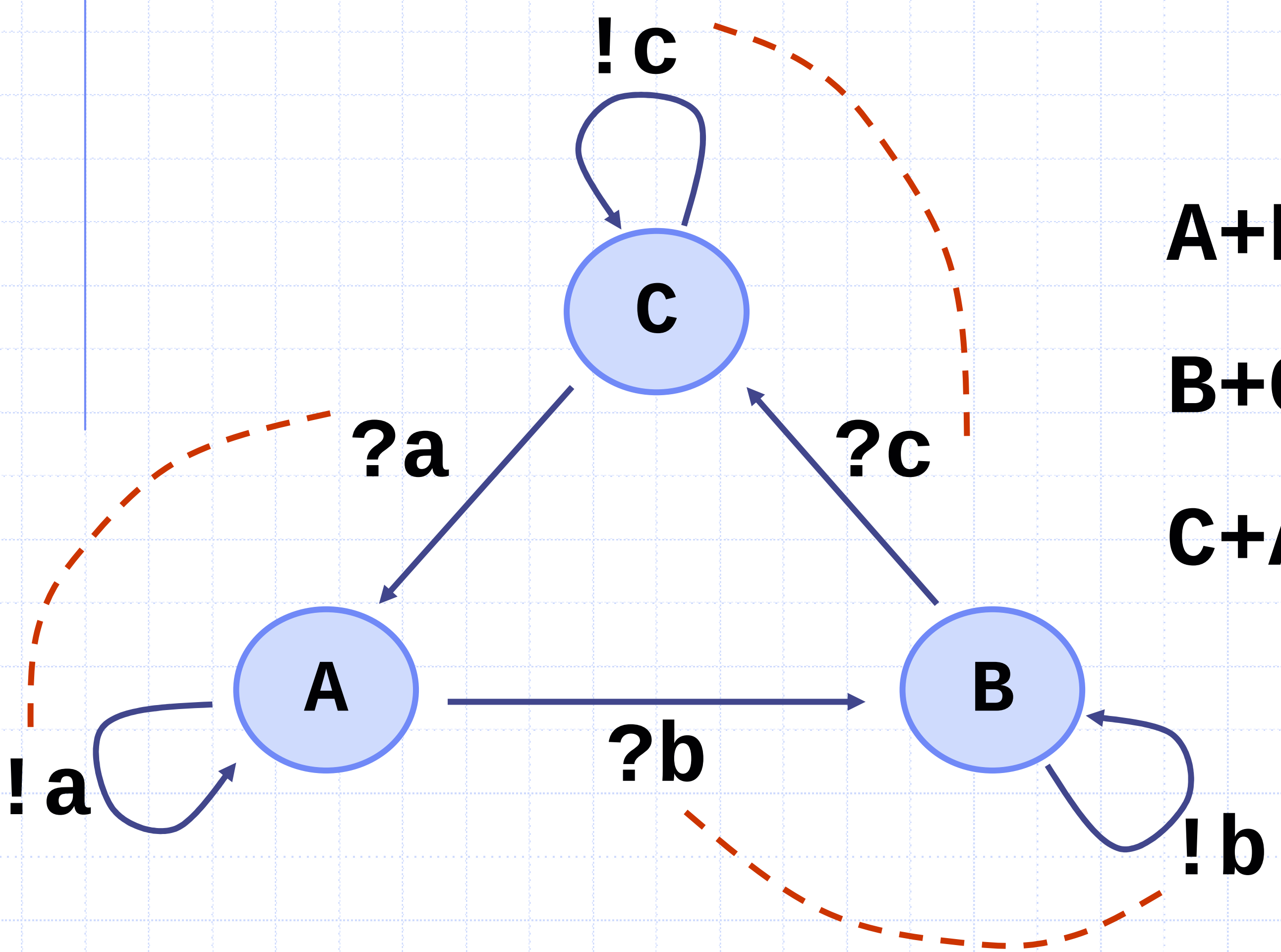
$a^n\tau b^n$

- Starting population: **A|A'**

CGF = Basic Chemistry [TCS08]



A nice example

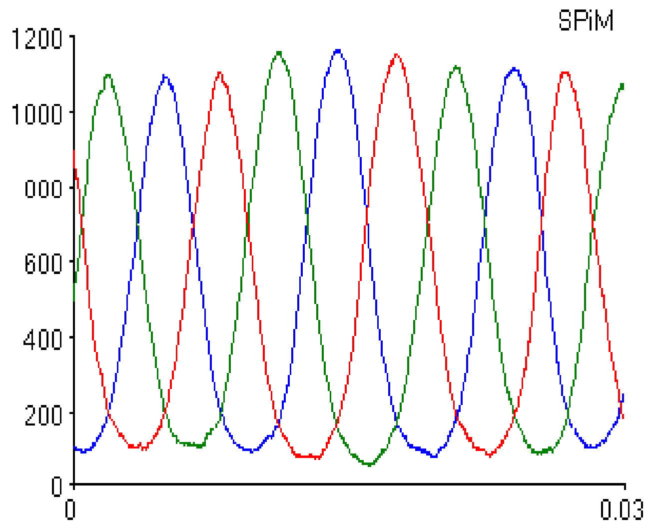


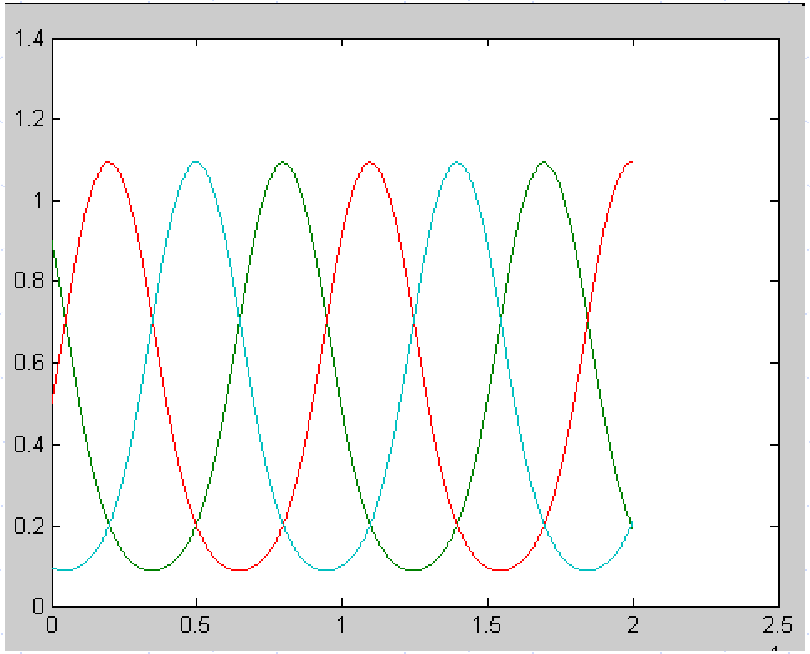
A+B → B+B

B+C → C+C

C+A → A+A

with a nice behaviour…





Discrete-State Semantics

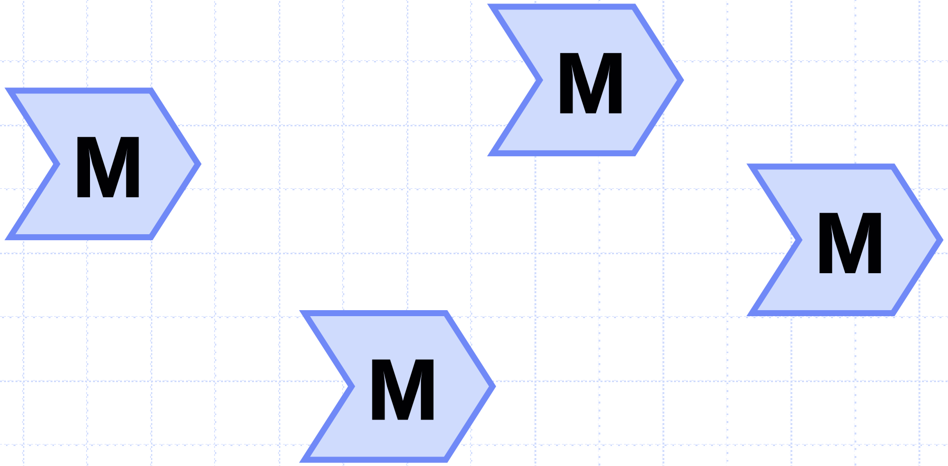
Continuous-State Semantics

Plan of the talk

- Basic Chemistry and Basic Biochemistry
  - Biochemistry = Chemistry + complexation
- Chemical Ground Form (CGF)
  - A process algebra for basic chemistry
- **Biochemical Ground Form (BGF)**
  - **A process algebra for basic biochemistry**
- Considered TERMINATION problems:
  - Existential termination in CGF (DECIDABLE)
  - Existential termination in BGF (UNDECIDIBLE)
  - Universal termination in CGF
    - Nondeterministic -all computations terminate- (DECIDABLE)
    - Probabilistic -terminate with probability 1- (UNDECIDABLE)

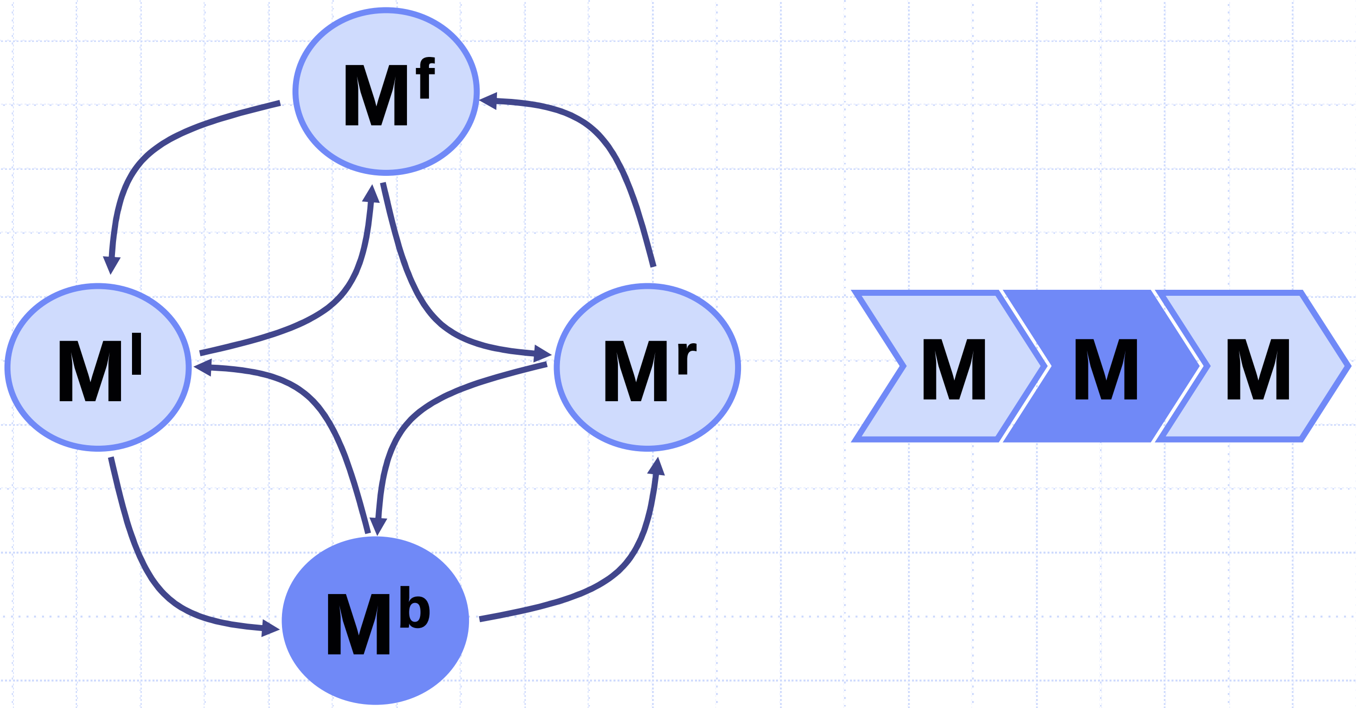
Polymerization

- Monomers associate and dissociate



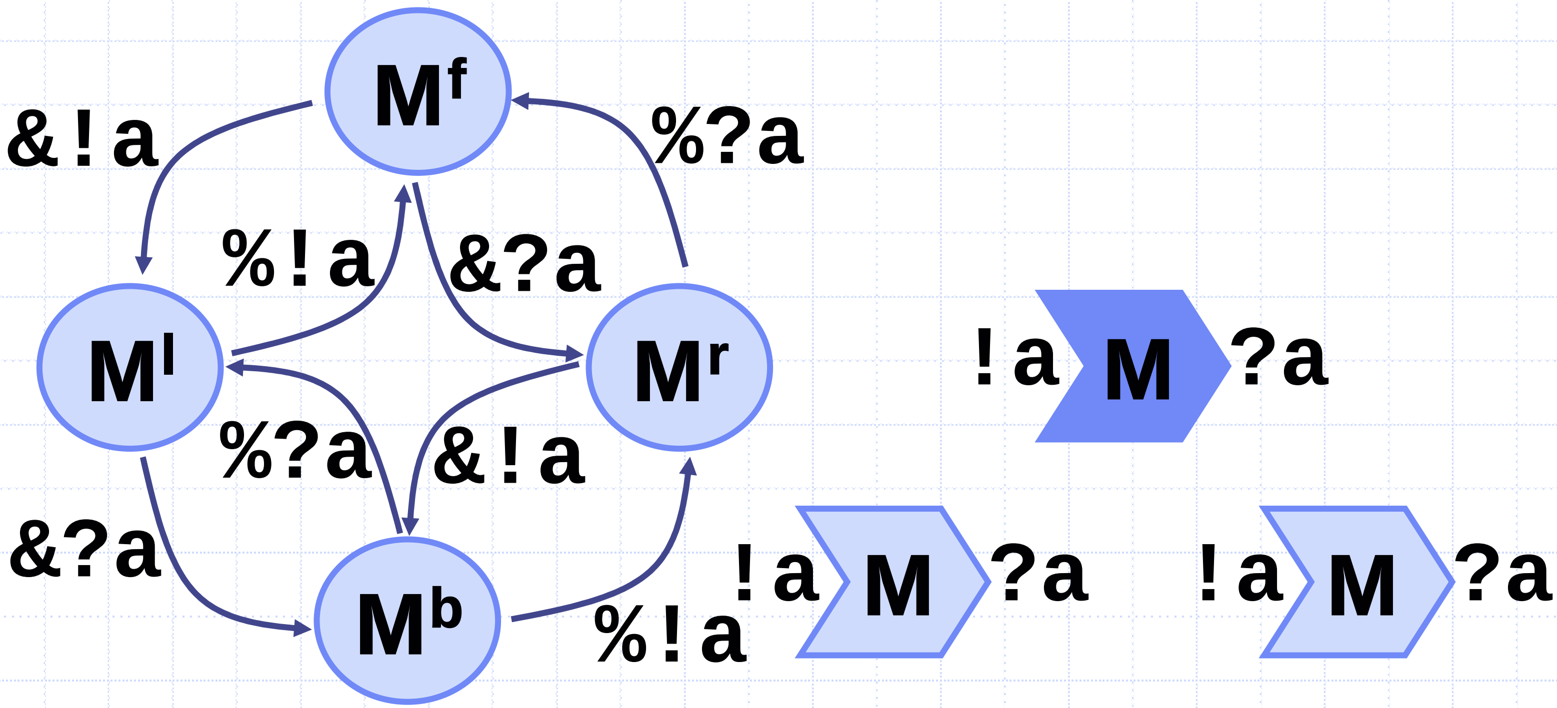
Association and Dissociation

- How to model the actin-like monomer behavior?



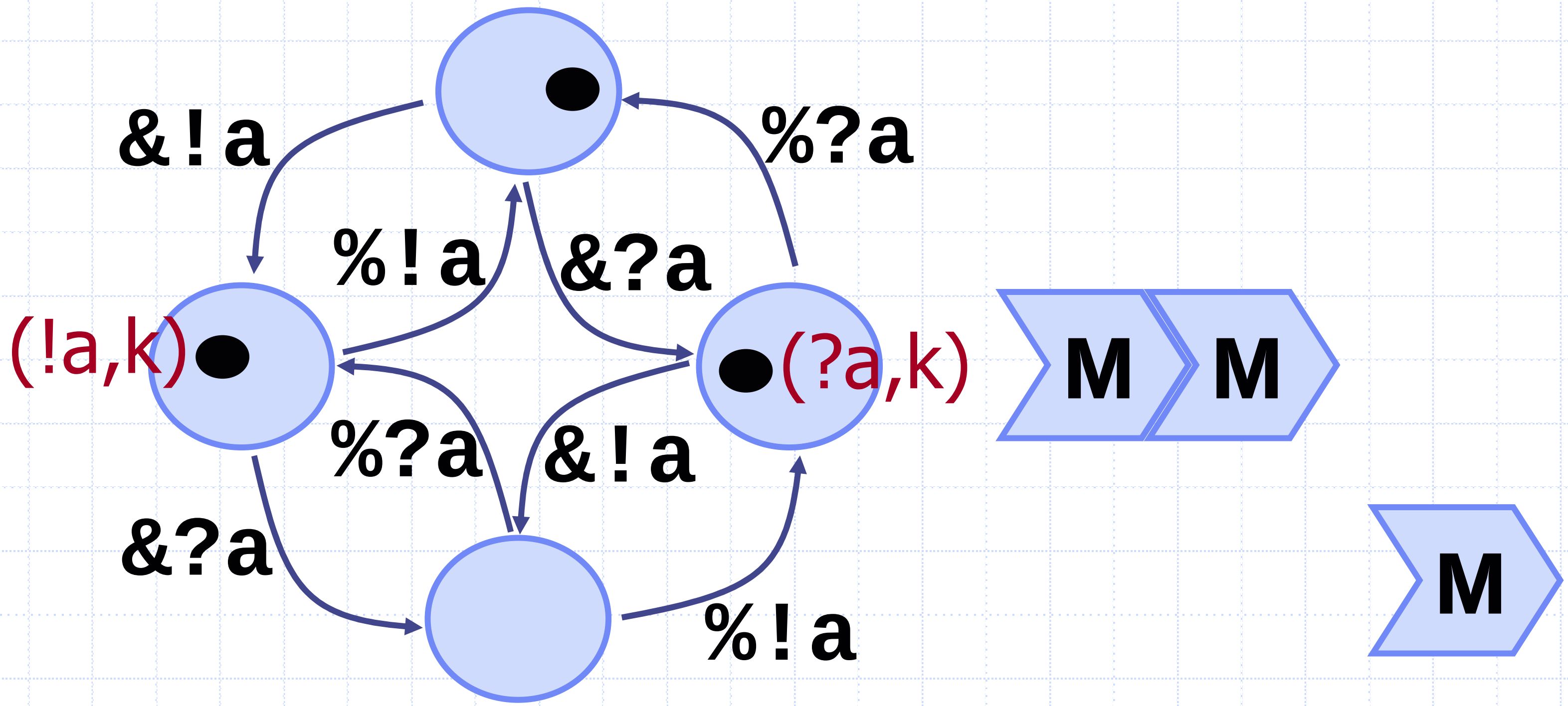
Association and Dissociation

- How to model the actin-like monomer behavior?



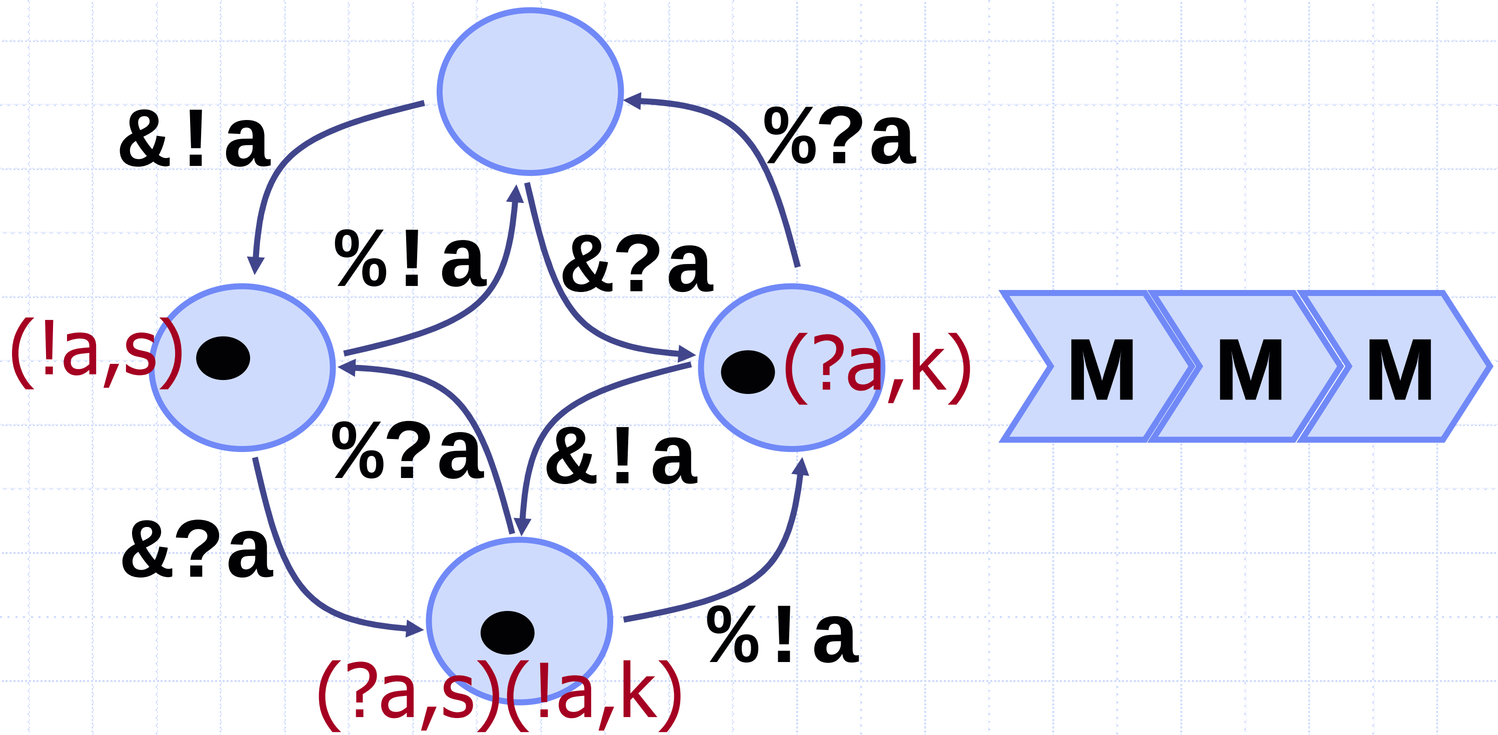
Association histories

- Each association has a unique key
  - Keys are stored in the molecule's history



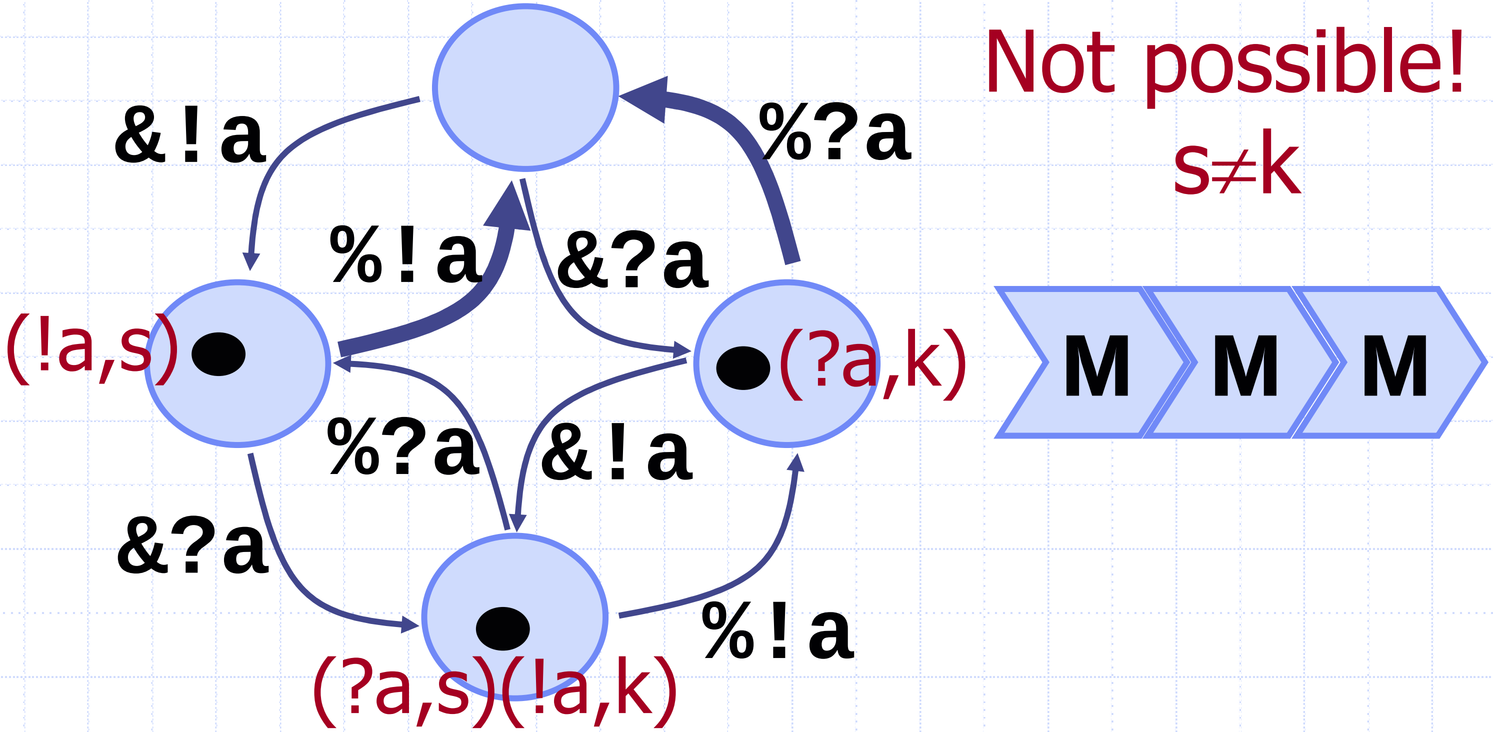
Association histories

- Each association has a unique key
  - Keys are stored in the molecule's history



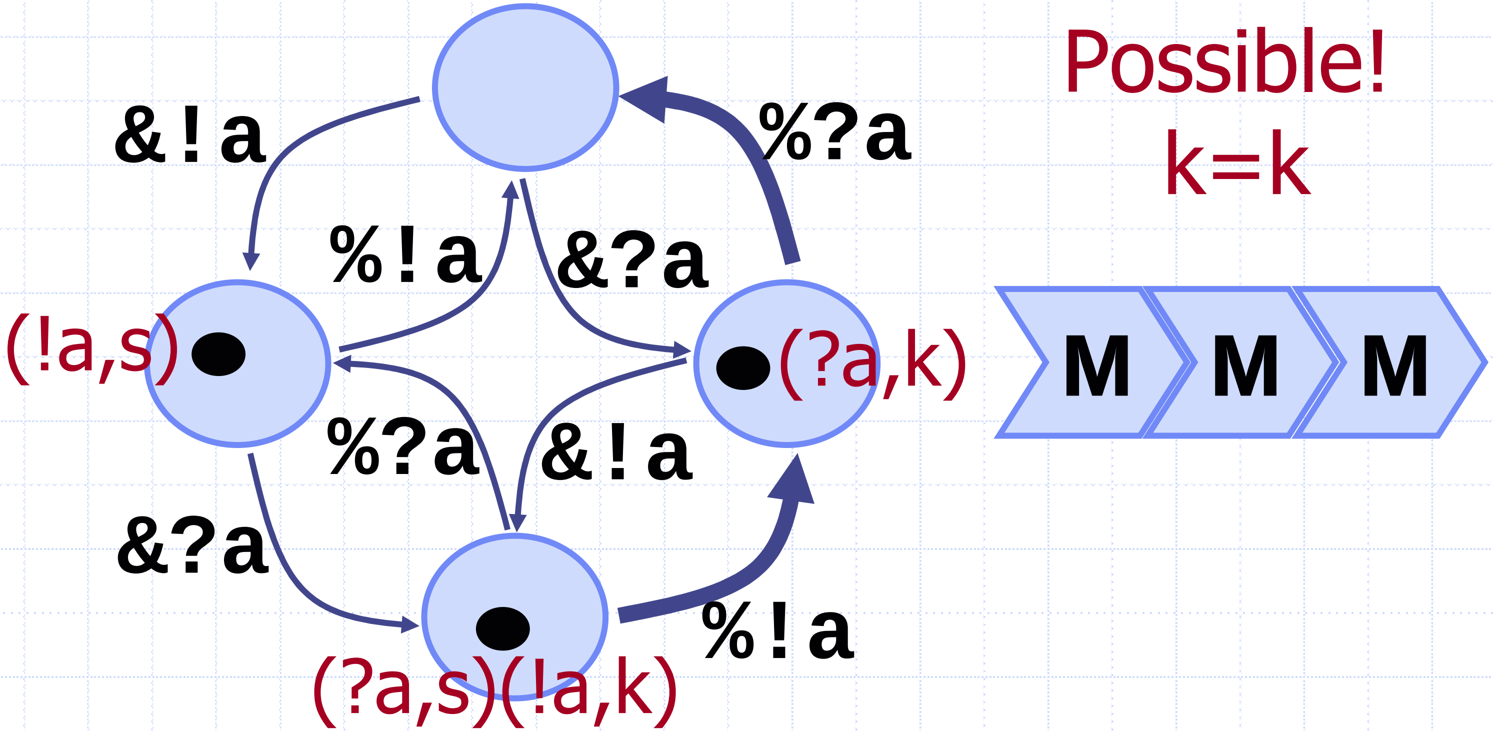
Association histories

- Each association has a unique key
  - Keys are stored in the molecule's history



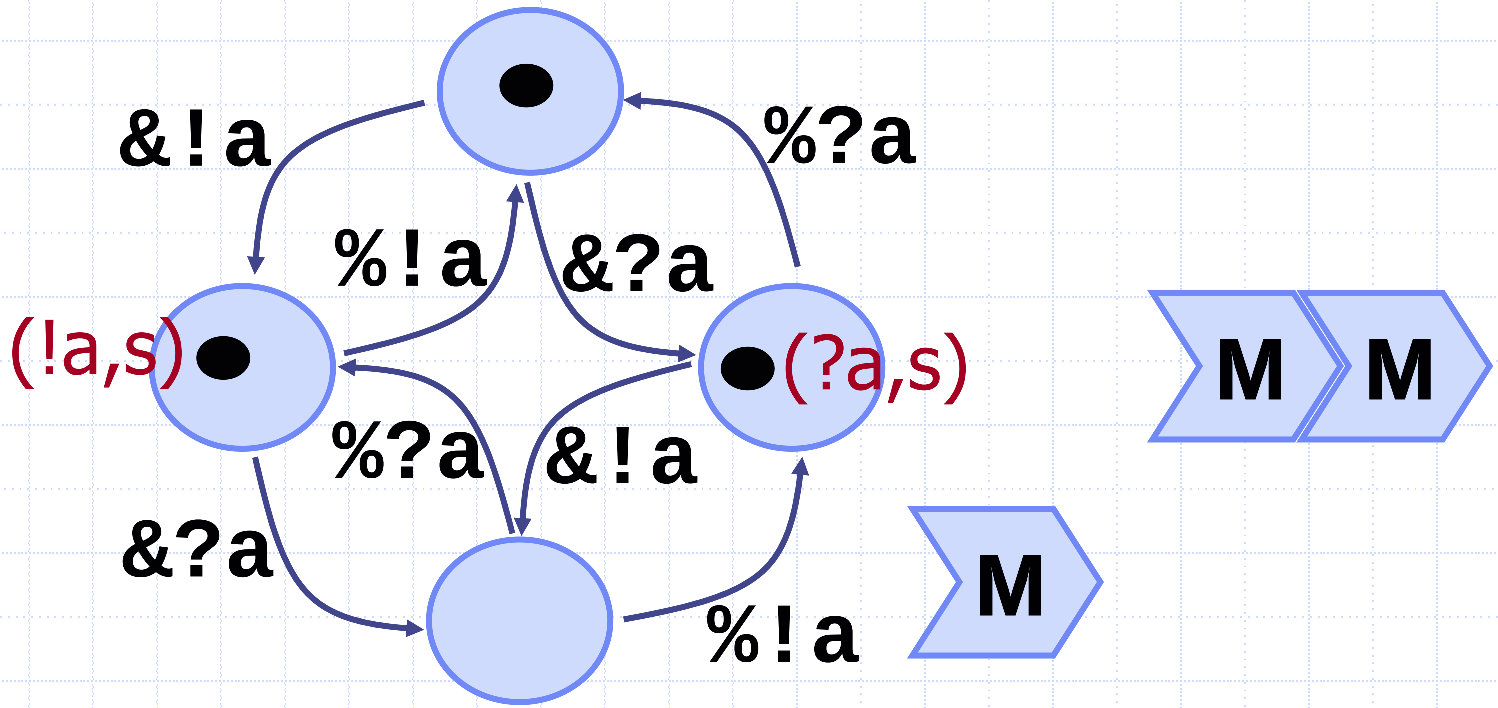
Association histories

- Each association has a unique key
  - Keys are stored in the molecule's history



Association histories

- Each association has a unique key
  - Keys are stored in the molecule's history



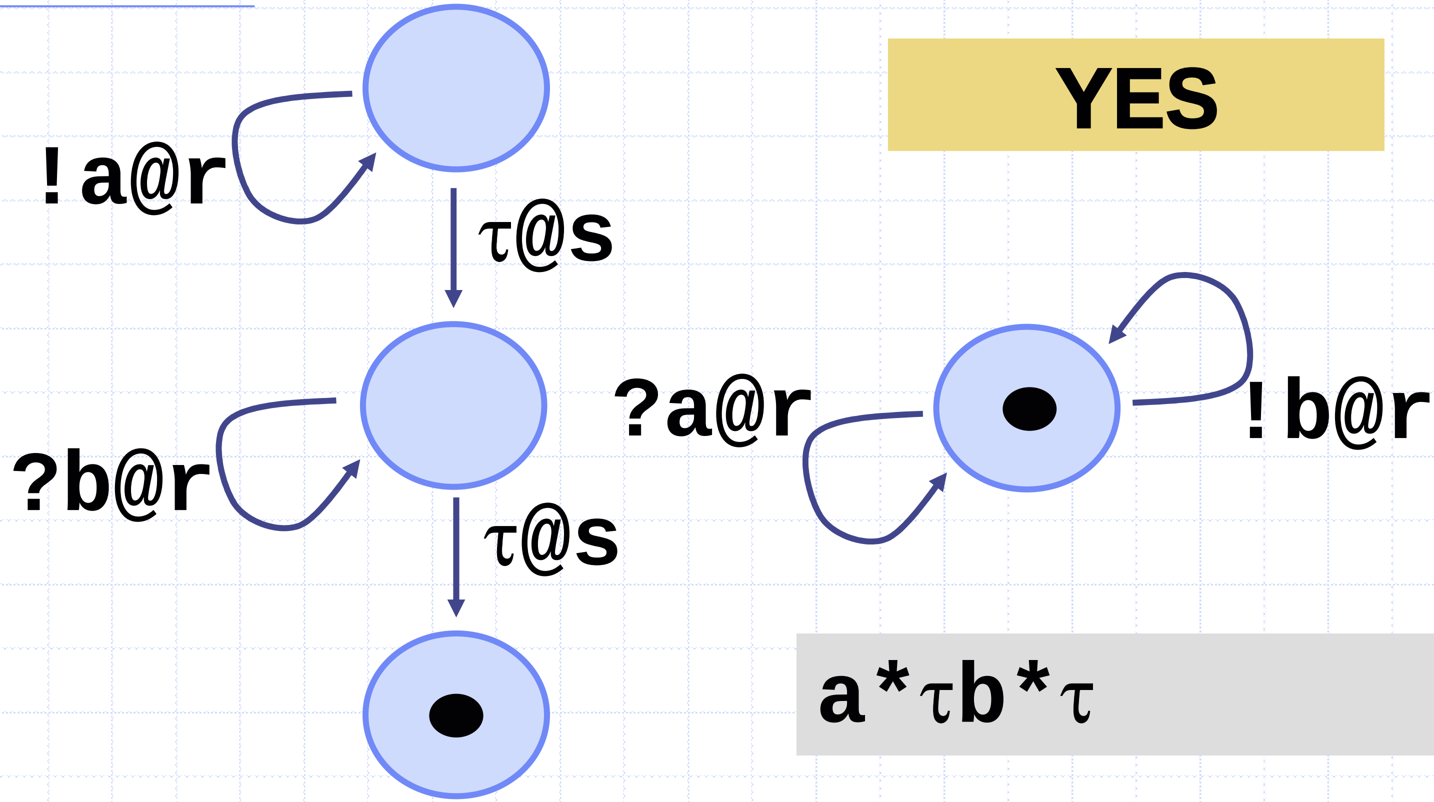
Plan of the talk

- Basic Chemistry and Basic Biochemistry
  - Biochemistry = Chemistry + complexation
- Chemical Ground Form (CGF)
  - A process algebra for basic chemistry
- Biochemical Ground Form (BGF)
  - A process algebra for basic biochemistry
- Considered TERMINATION problems:
  - **Existential termination in CGF (DECIDABLE)**
  - Existential termination in BGF (UNDECIDIBLE)
  - Universal termination in CGF
    - Nondeterministic -all computations terminate- (DECIDABLE)
    - Probabilistic -terminate with probability 1- (UNDECIDABLE)

Existential termination for CGF

- Given a CGF system, decide whether there exists a computation leading to a deadlock

Example 1: does it terminate?

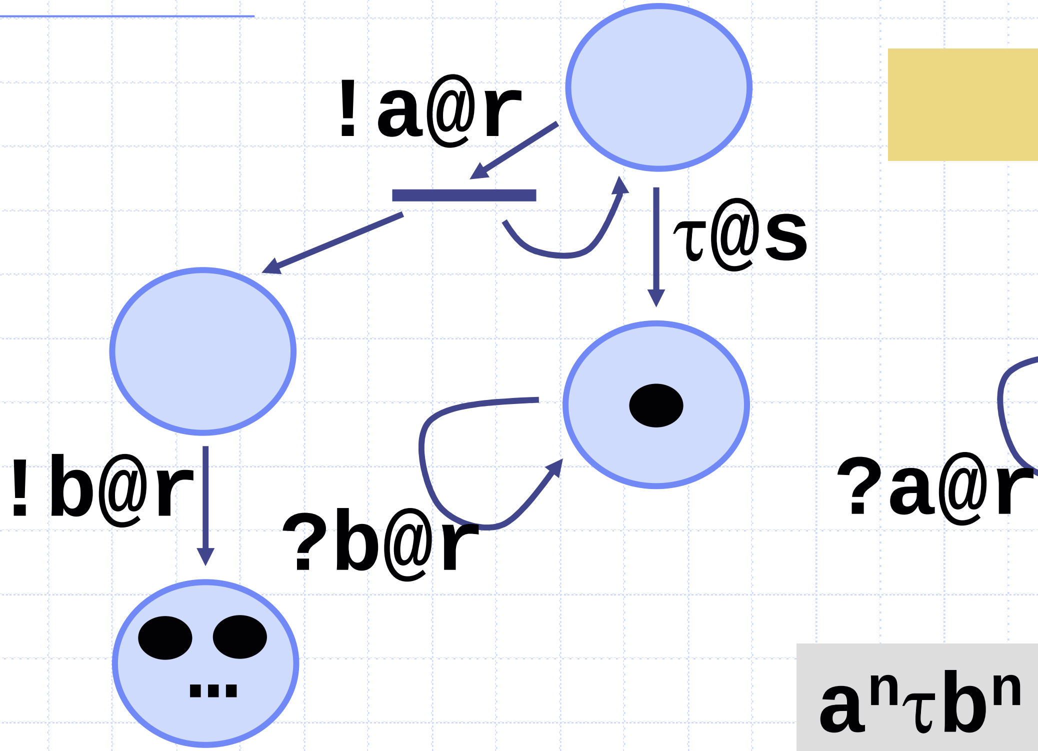


YES

$a^*\tau b^*\tau$

- Starting population: **A|A'**

Example 2: does it terminate?

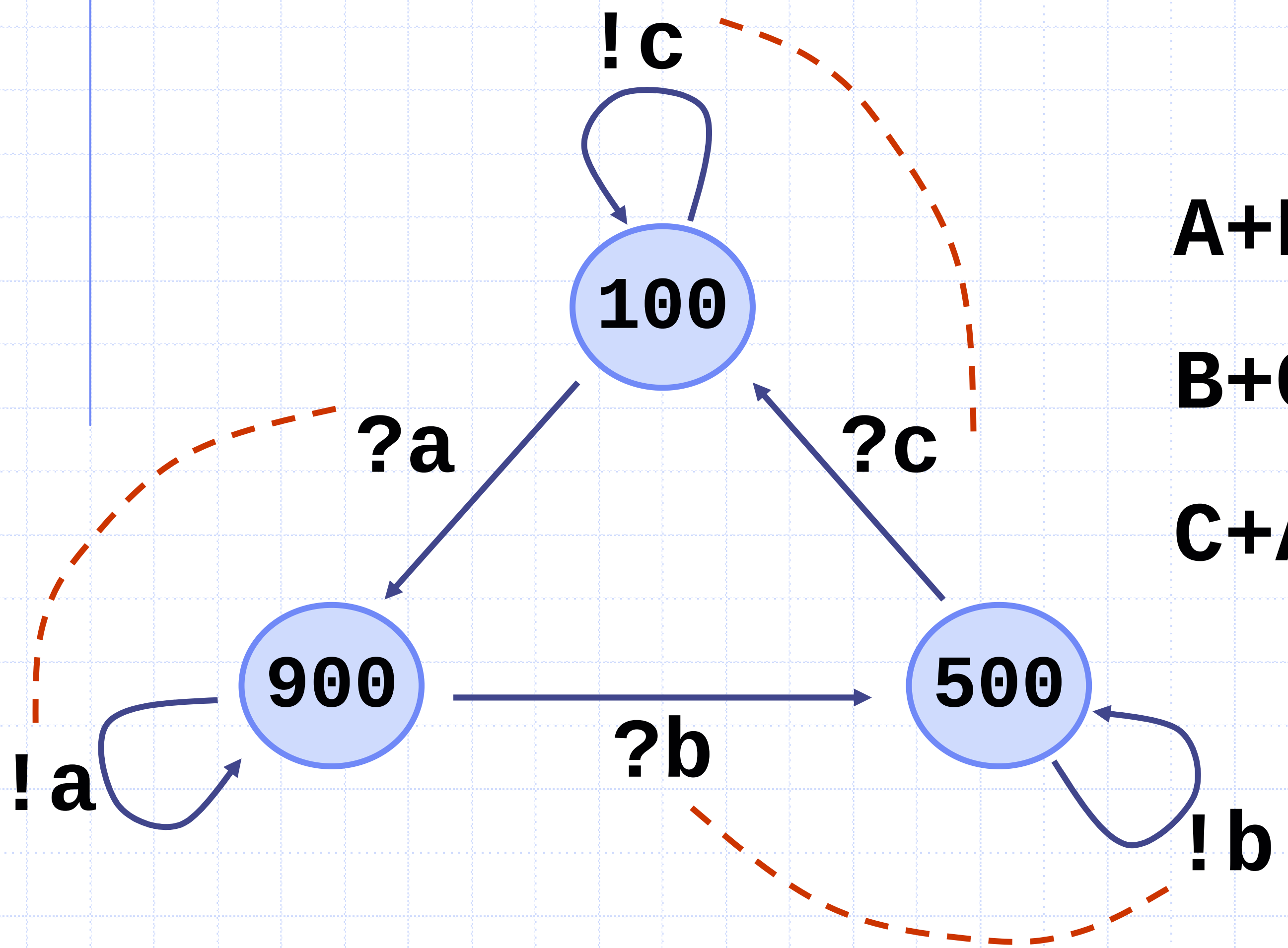


YES

$a^n \tau b^n$

- Starting population: **A|A′**

Example 3: does it terminate?

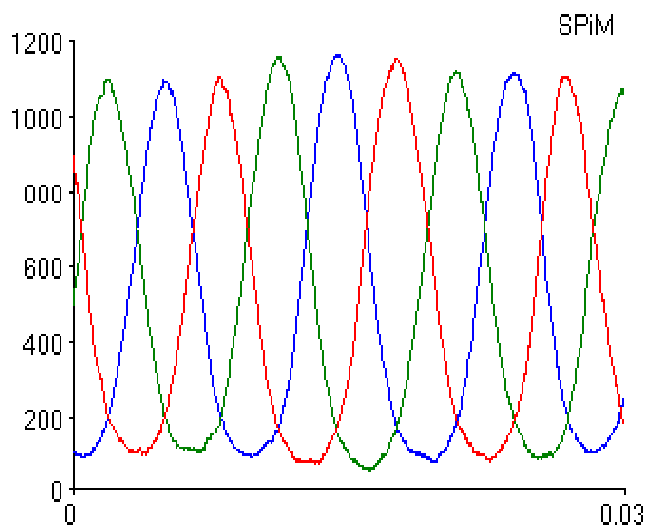


A+B → B+B

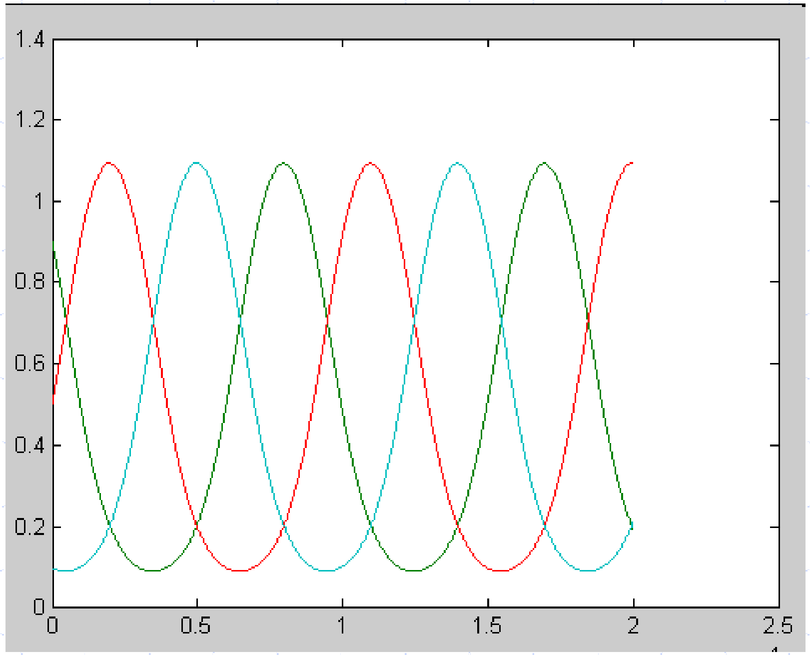
B+C → C+C

C+A → A+A

with a nice behaviour…

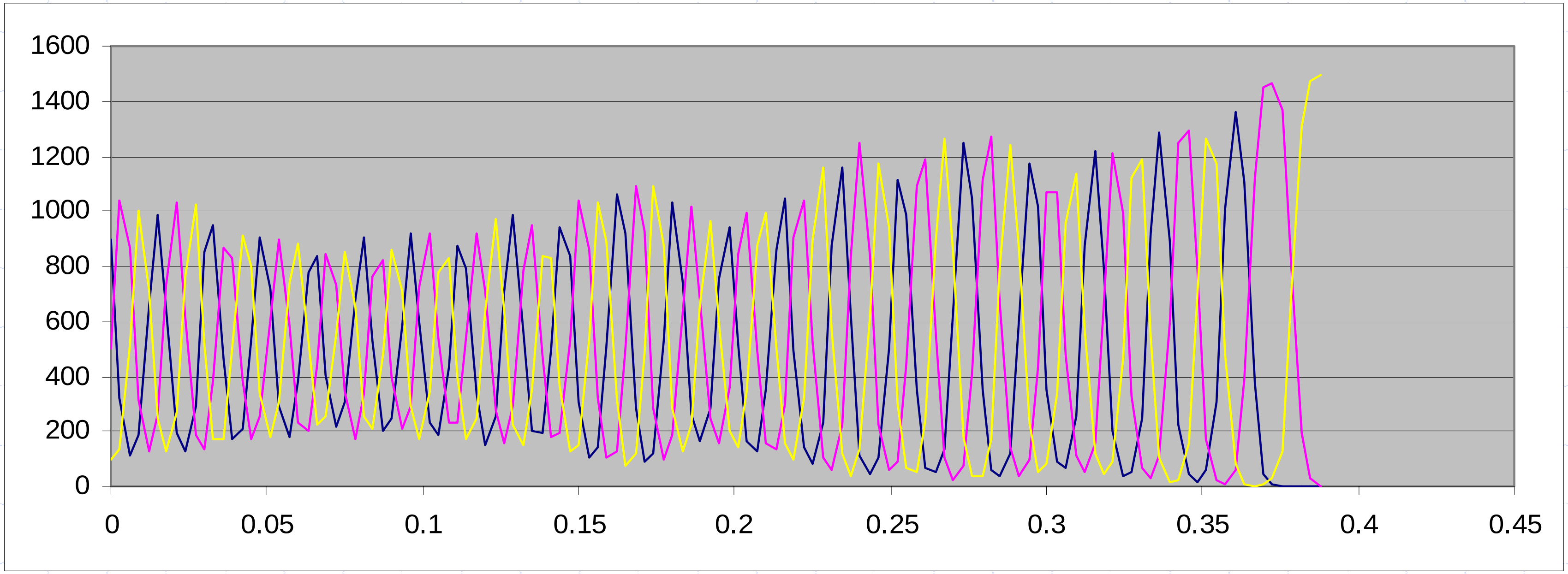


Discrete-State Semantics



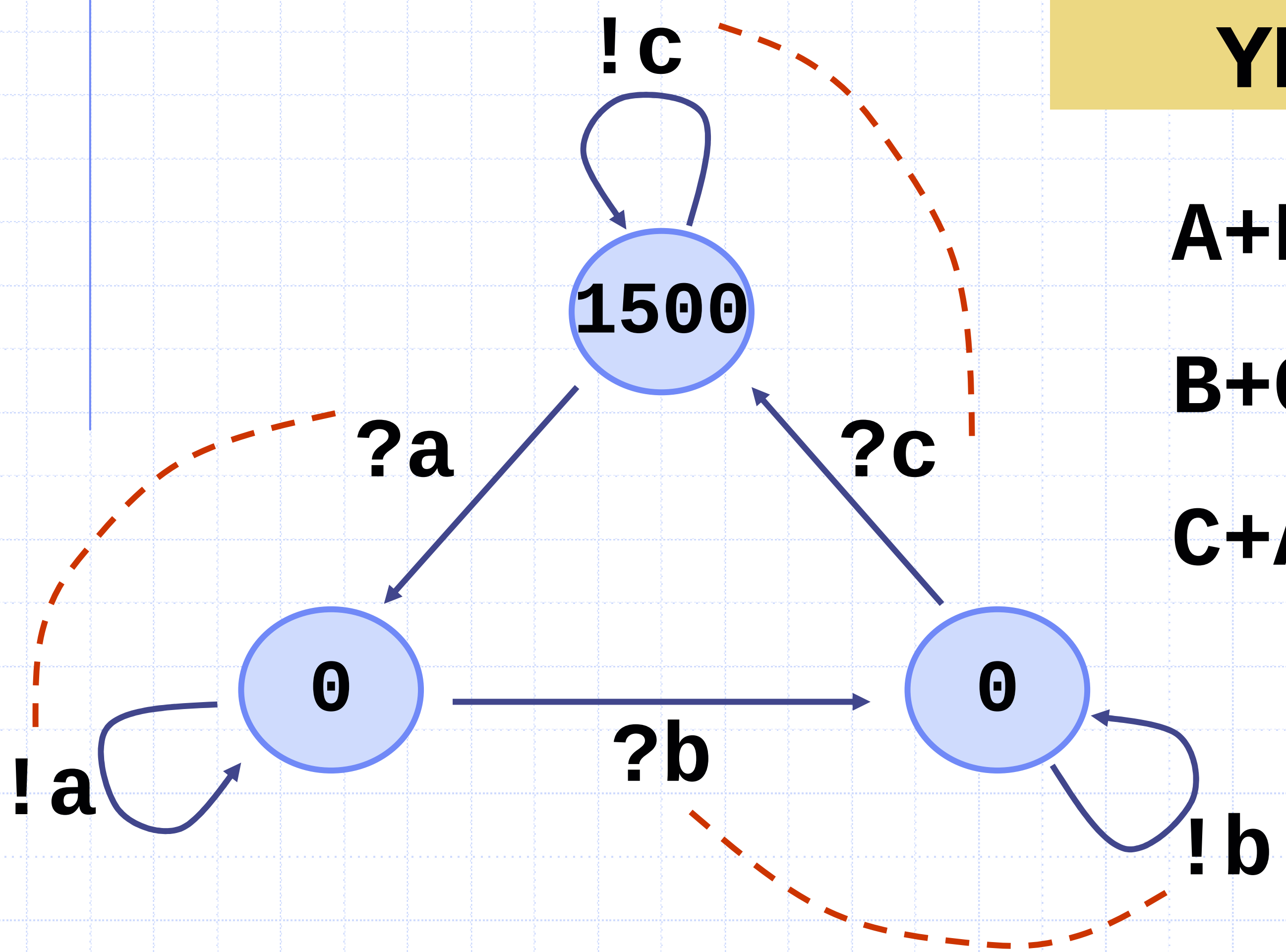
Continuous-State Semantics

with a nice behaviour…



But in a longer simulation…

Example 3: does it terminate?



YES

A+B → B+B

B+C → C+C

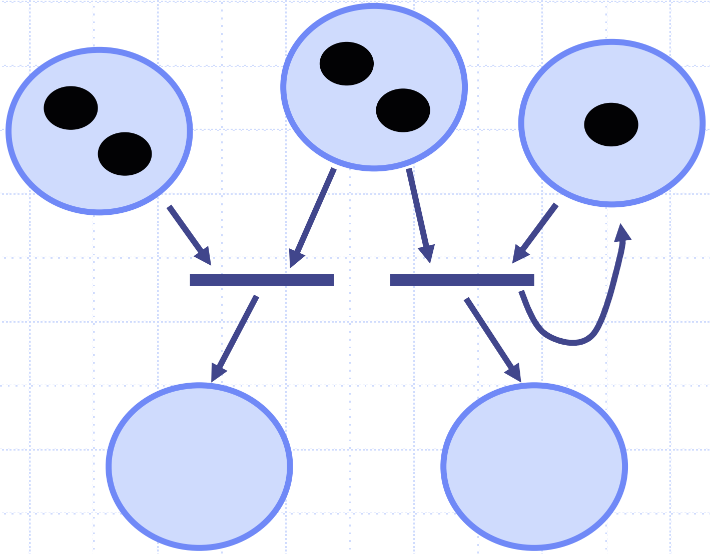
C+A → A+A

Decidability of termination

- We reduce existential termination for CGF to termination for Petri Nets
  - Petri Nets is an interesting infinite state system in which many properties (reachability, coverability, termination, divergence,…) are decidable

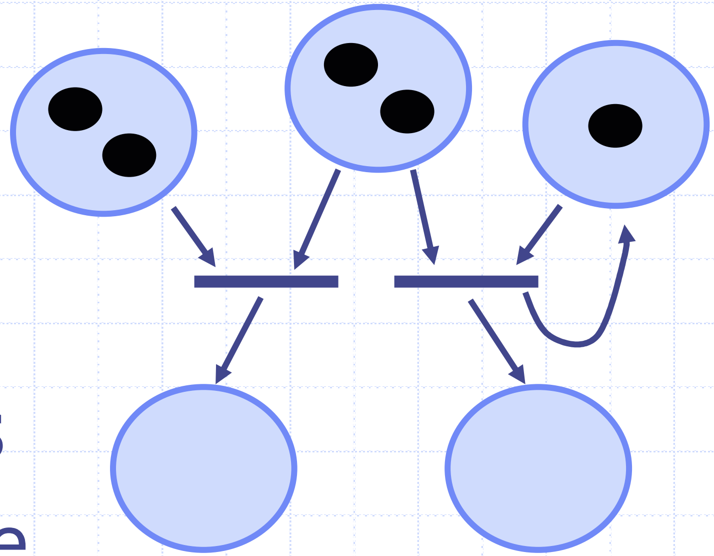
Petri nets

- A Petri net is a triple
  - A finite set of **Places**
  - A finite set of **Transitions**: pairs of multisets of places (preset,postset)
  - An initial **marking** (multiset of places)



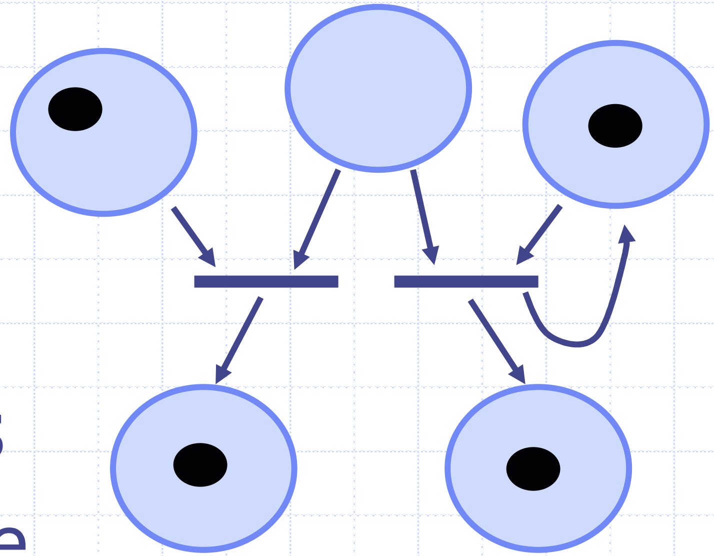
Petri nets

- A transition is enabled
  - when it is possible to consume tokens in the preset
- When a transition fires
  - tokens are placed in the postset



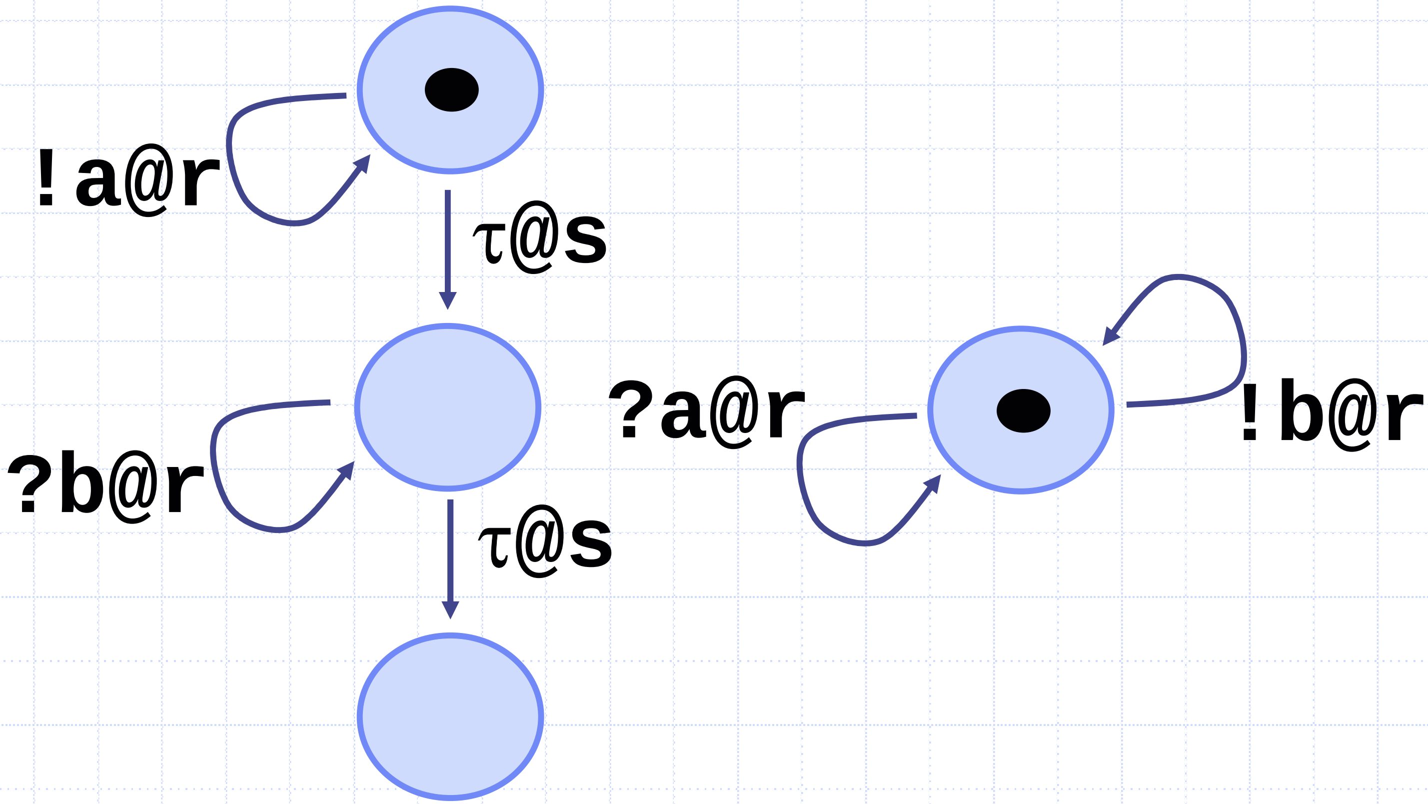
Petri nets

- A transition is enabled
  - when it is possible to consume tokens in the preset
- When a transition fires
  - tokens are placed in the postset



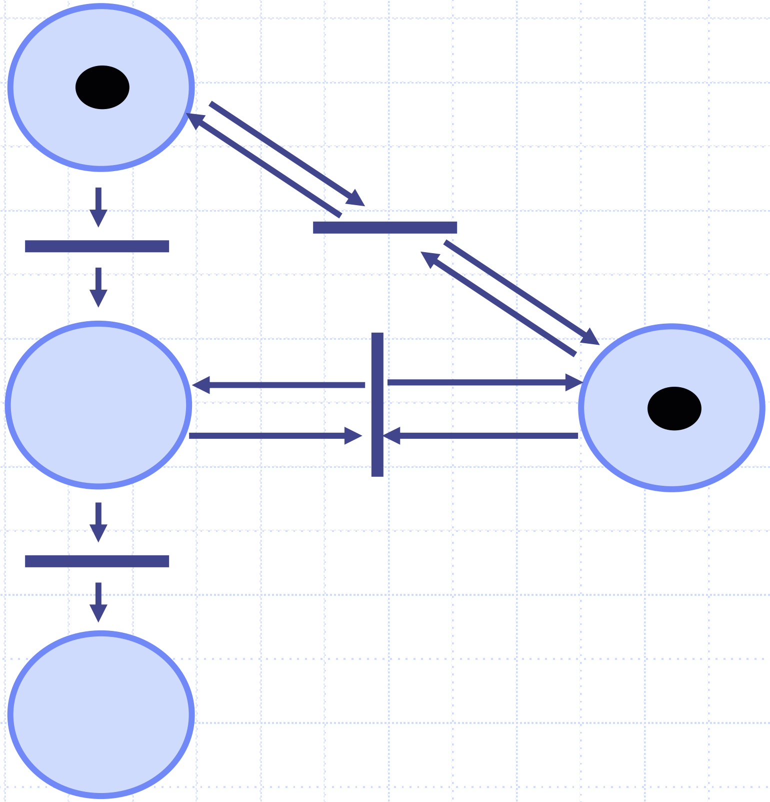
A Petri net semantics for CGF

- One place for each Species
- One transition for each reaction



A Petri net semantics for CGF

- One place for each Species
- One transition for each reaction



Plan of the talk

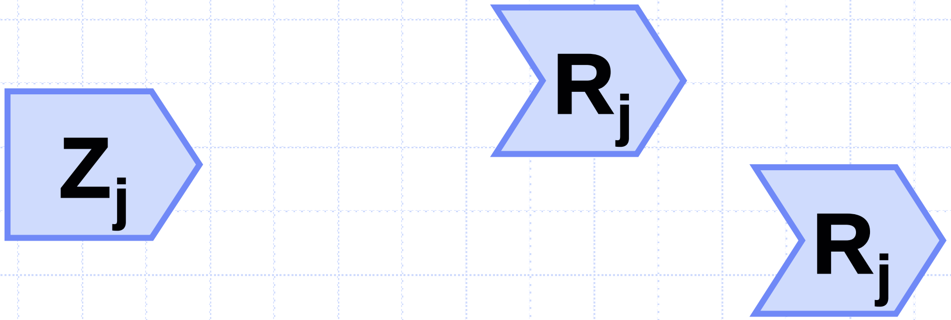
- Basic Chemistry and Basic Biochemistry
  - Biochemistry = Chemistry + complexation
- Chemical Ground Form (CGF)
  - A process algebra for basic chemistry
- Biochemical Ground Form (BGF)
  - A process algebra for basic biochemistry
- Considered TERMINATION problems:
  - Existential termination in CGF (DECIDABLE)
  - **Existential termination in BGF (UNDECIDIBLE)**
  - Universal termination in CGF
    - Nondeterministic -all computations terminate- (DECIDABLE)
    - Probabilistic -terminate with probability 1- (UNDECIDABLE)

Turing completeness of BGF

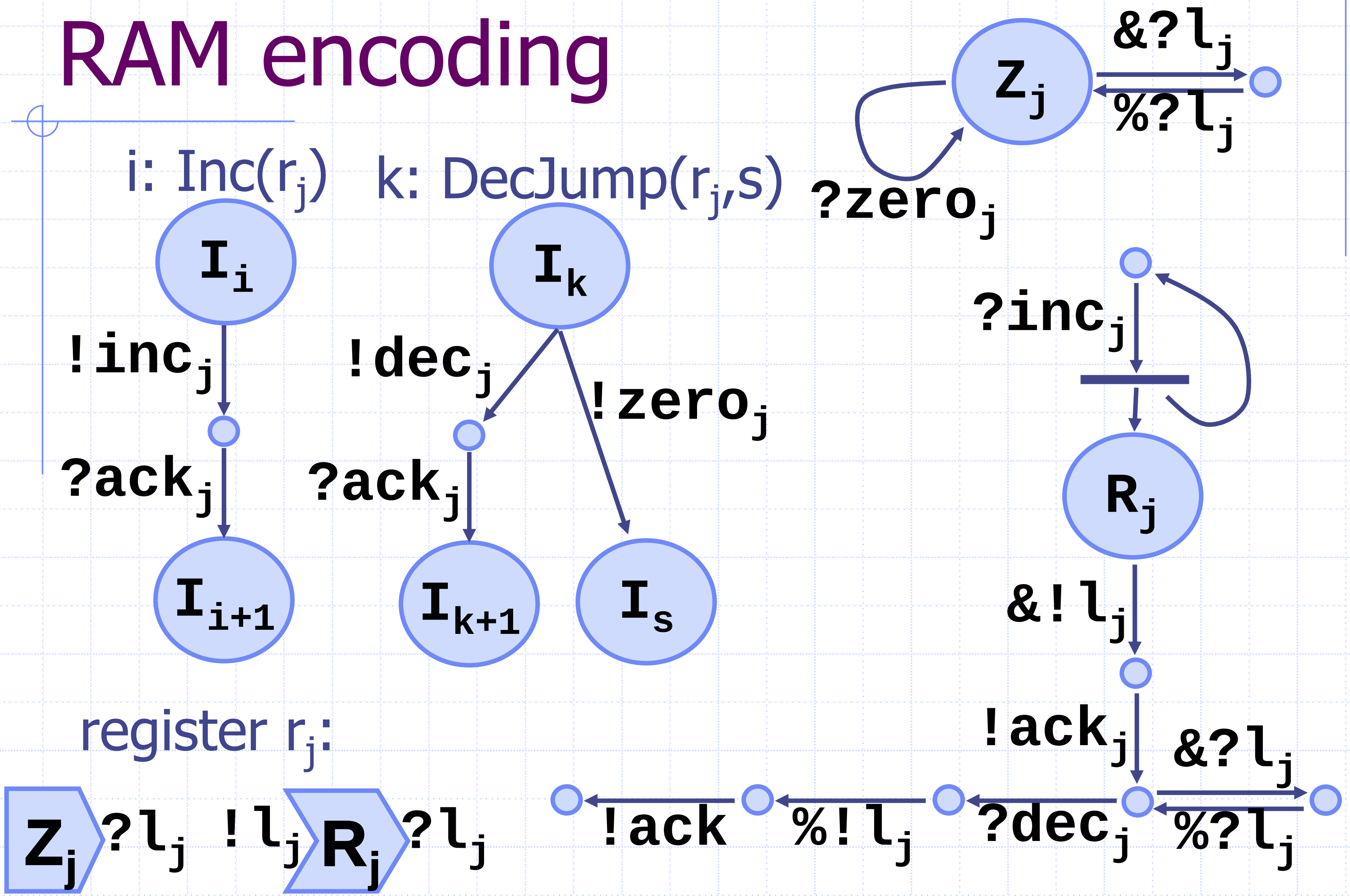
- In BGF we model Random Access Machines: [Min67]
  - **Registers**: $r_1 \dots r_n$ hold natural numbers
  - **Program**: sequence of numbered instructions
    - **i: Inc(r_j)**: add 1 to the content of r_j and go to the next instruction
    - **i: DecJump(r_j,s)**: if the content of r_j is not 0 then decrease by 1 and go to the next instruction; otherwise jump to instruction s

Registers as Linearly growing polymer

- Initially empty register r_j: a seed $\mathbf{Z_j}$
- Increment on r_j: produce a new monomer and associate it to the polymer
- Decrement on r_j: remove last monomer



RAM encoding



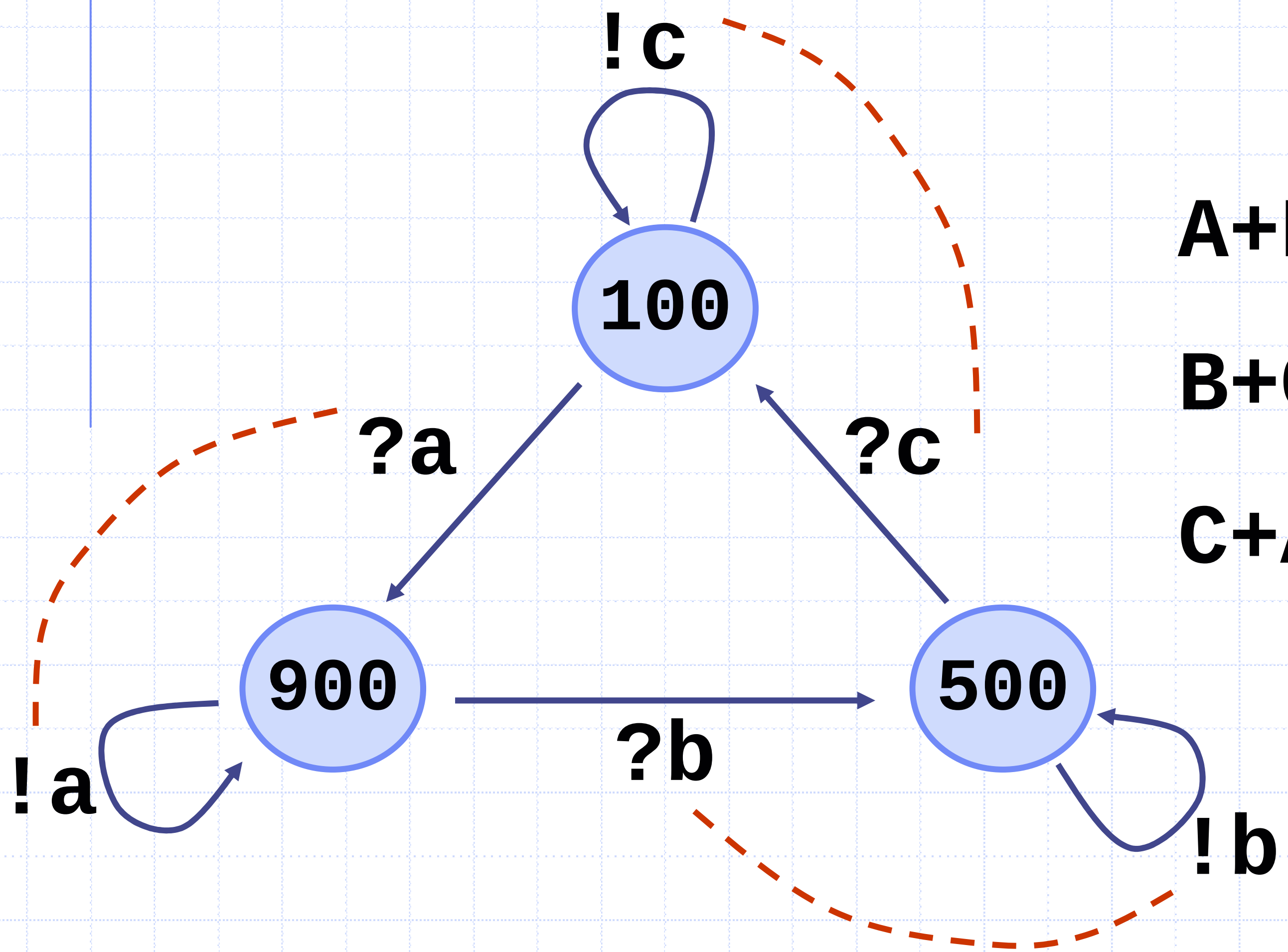
Plan of the talk

- Basic Chemistry and Basic Biochemistry
  - Biochemistry = Chemistry + complexation
- Chemical Ground Form (CGF)
  - A process algebra for basic chemistry
- Biochemical Ground Form (BGF)
  - A process algebra for basic biochemistry
- Considered TERMINATION problems:
  - Existential termination in CGF (DECIDABLE)
  - Existential termination in BGF (UNDECIDIBLE)
  - Universal termination in CGF
    - **Nondeterministic -all computations terminate- (DECIDABLE)**
    - Probabilistic -terminate with probability 1- (UNDECIDABLE)

Petri Nets strike back…

- In Petri nets, termination of all computations is decidable
  - the translation from CGF to Petri nets allows us to prove that (nondeterministic) universal termination in CGF is decidable

Example 3: does it (nondeterministically) universally terminate?

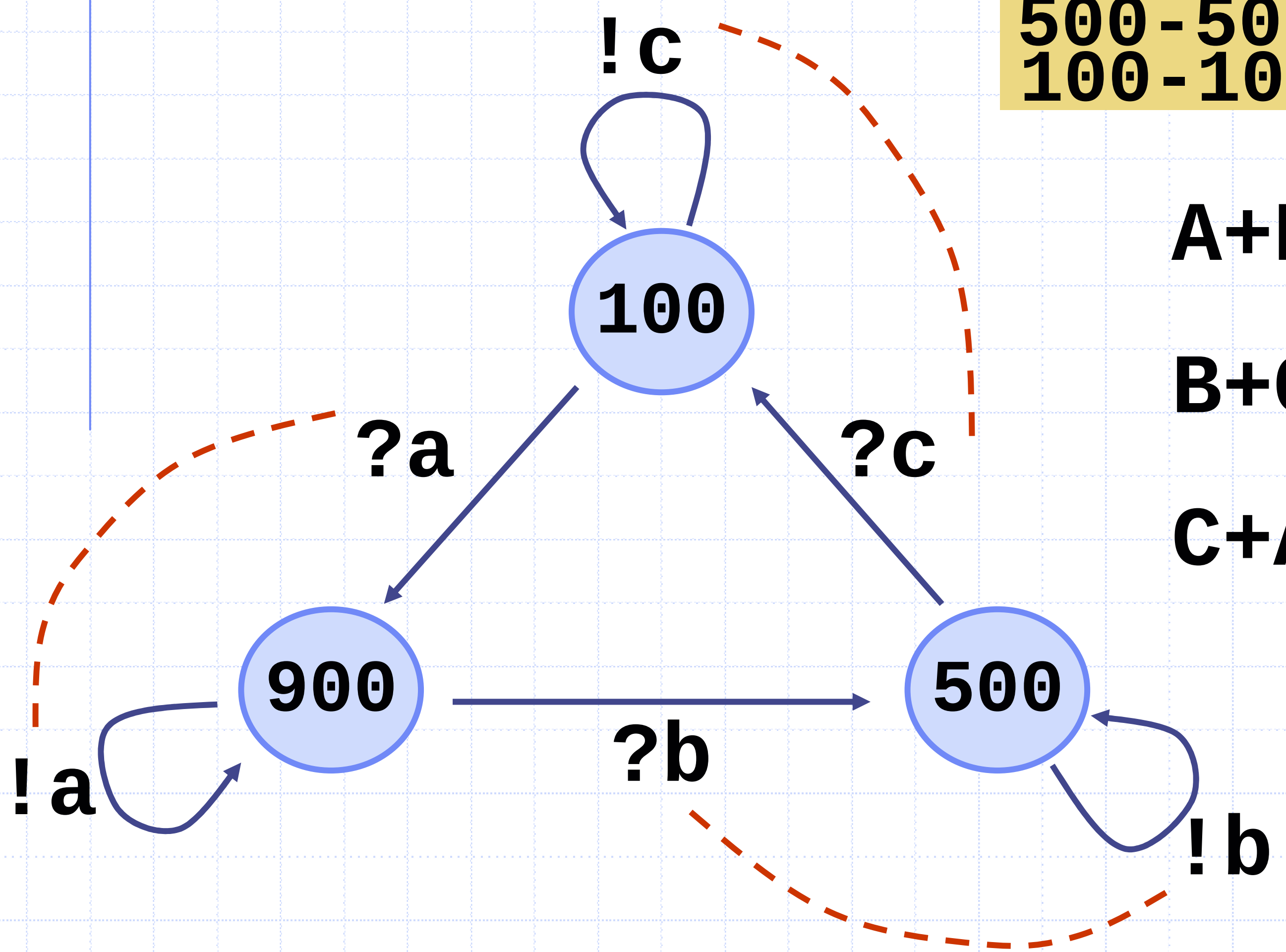


A+B → B+B

B+C → C+C

C+A → A+A

Example 3: does it (nondeterministically) universally terminate?



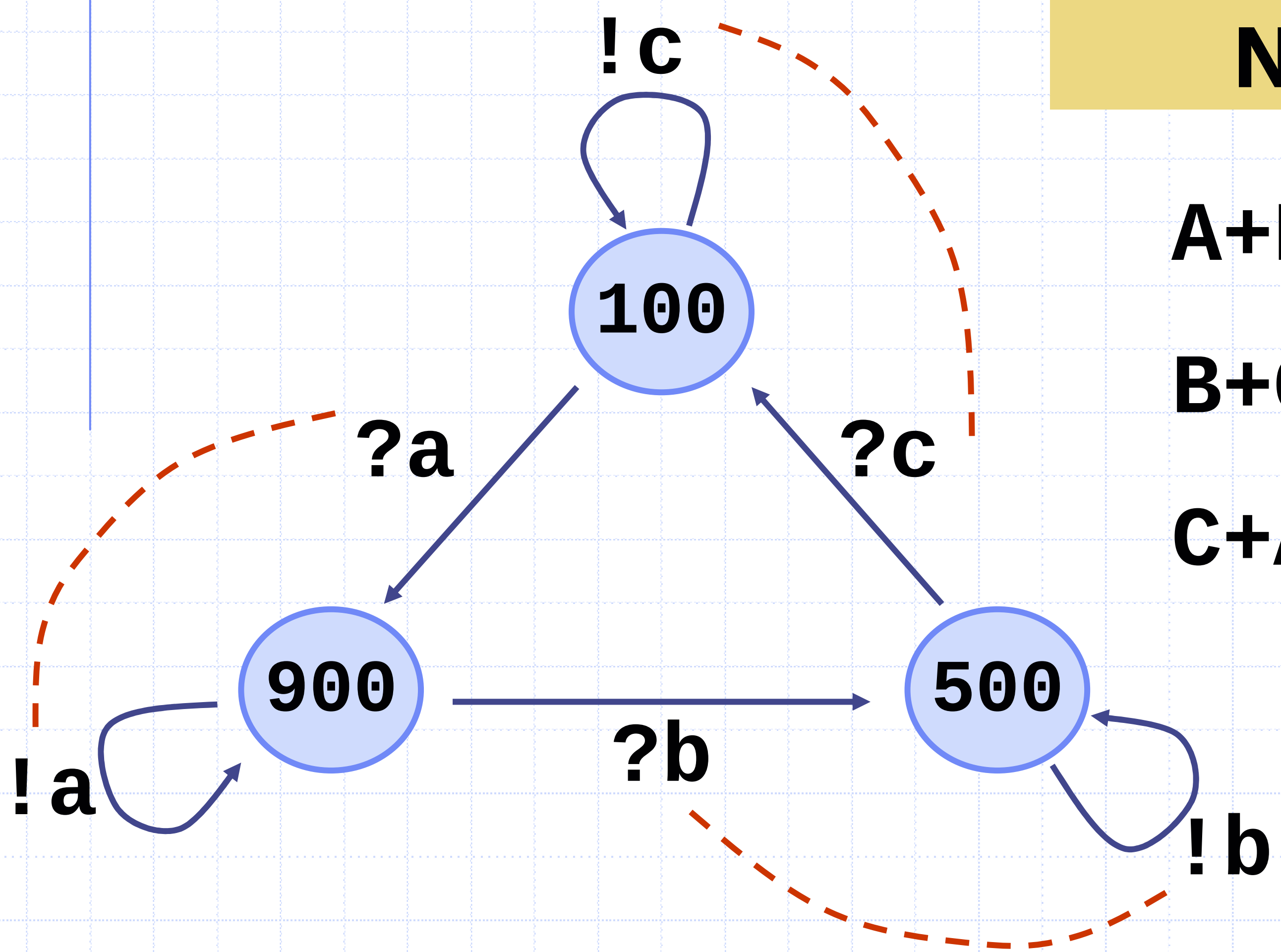
900-899-899-900..
500-501-500-500..
100-100-101-100..

A+B → B+B

B+C → C+C

C+A → A+A

Example 3: does it (nondeterministically) universally terminate?



NO

A+B → B+B

B+C → C+C

C+A → A+A

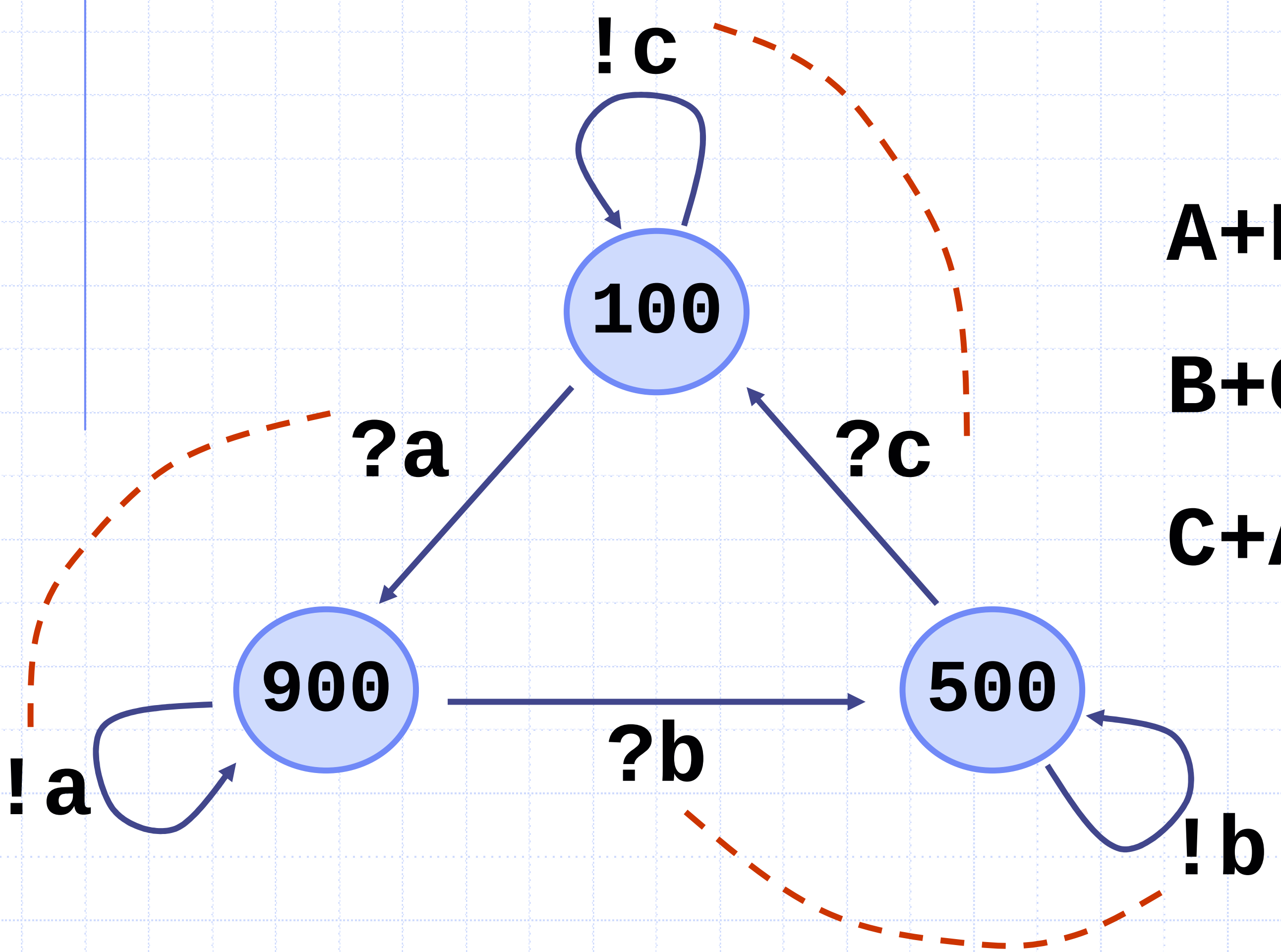
Plan of the talk

- Basic Chemistry and Basic Biochemistry
  - Biochemistry = Chemistry + complexation
- Chemical Ground Form (CGF)
  - A process algebra for basic chemistry
- Biochemical Ground Form (BGF)
  - A process algebra for basic biochemistry
- Considered TERMINATION problems:
  - Existential termination in CGF (DECIDABLE)
  - Existential termination in BGF (UNDECIDIBLE)
  - Universal termination in CGF
    - Nondeterministic -all computations terminate- (DECIDABLE)
    - **Probabilistic -terminate with probability 1- (UNDECIDABLE)**

Probabilistic universal termination

- Given a CGF system, decide whether the probability for the system to terminate is 1
  - This corresponds to checking whether there exists an infinite computation with associated probability > 0

Example 3: does it (probabilistically) universally terminate?

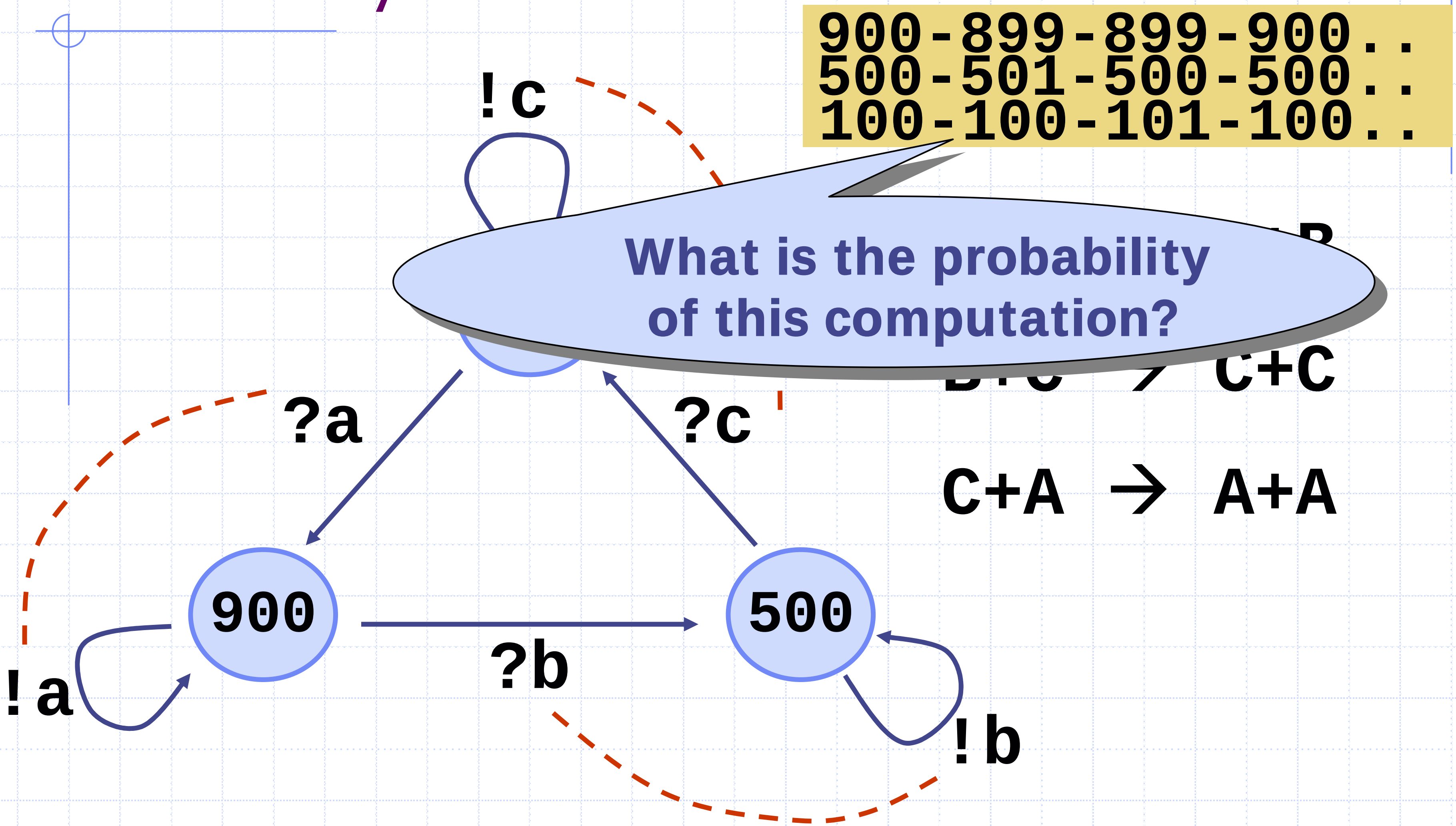


A+B → B+B

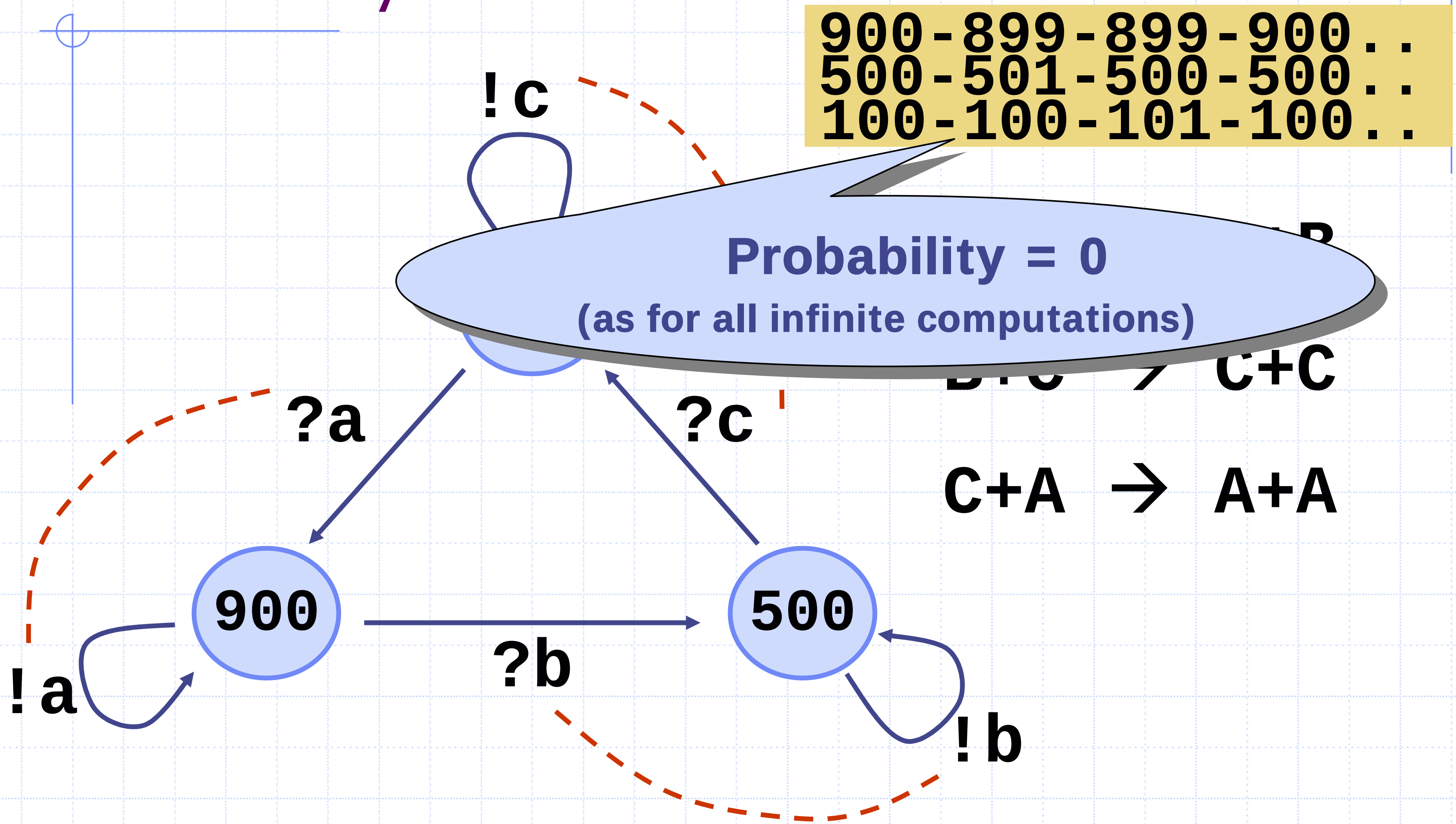
B+C → C+C

C+A → A+A

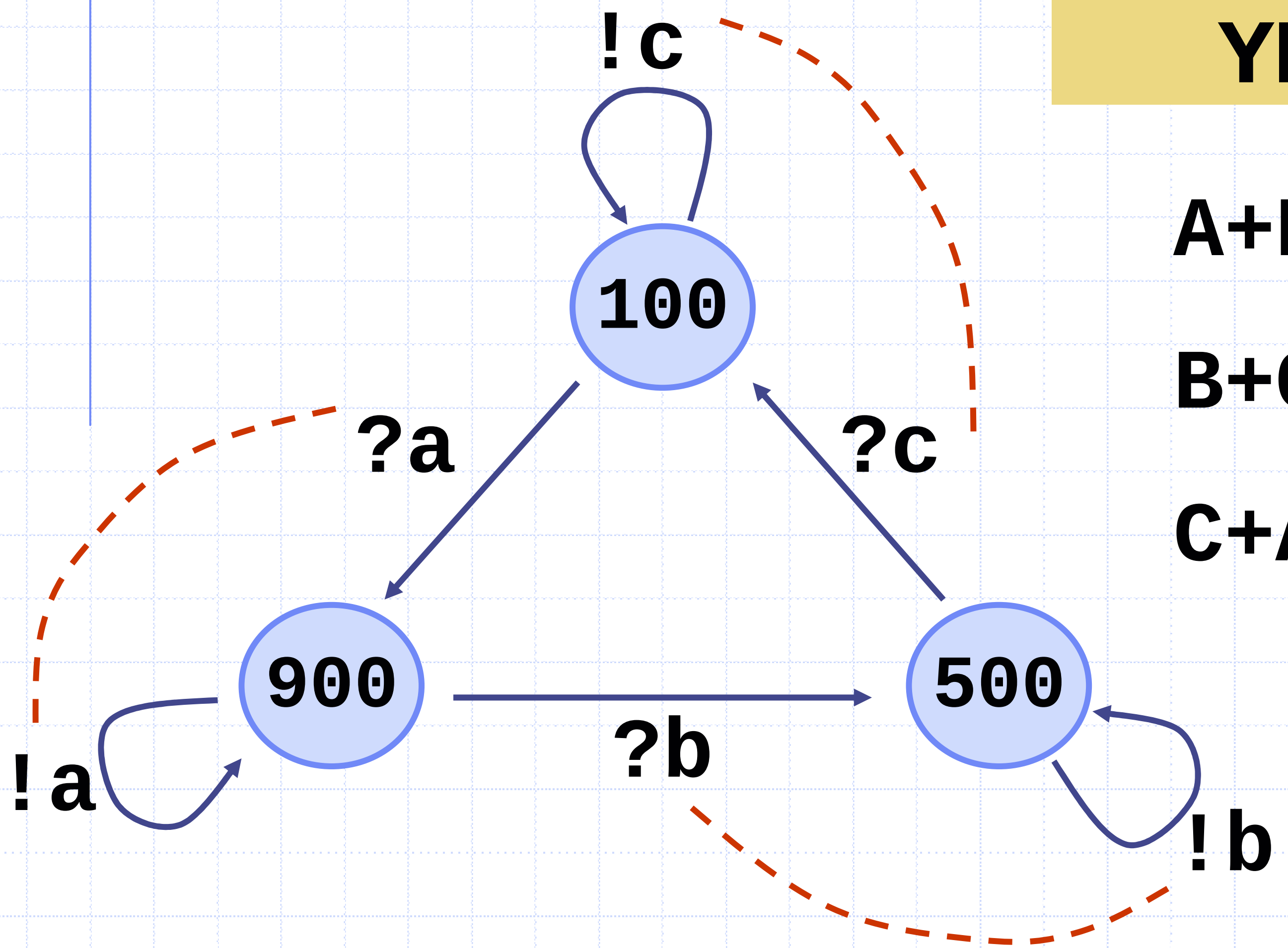
Example 3: does it (probabilistically) universally terminate?



Example 3: does it (probabilistically) universally terminate?



Example 3: does it (probabilistically) universally terminate?



YES

A+B → B+B

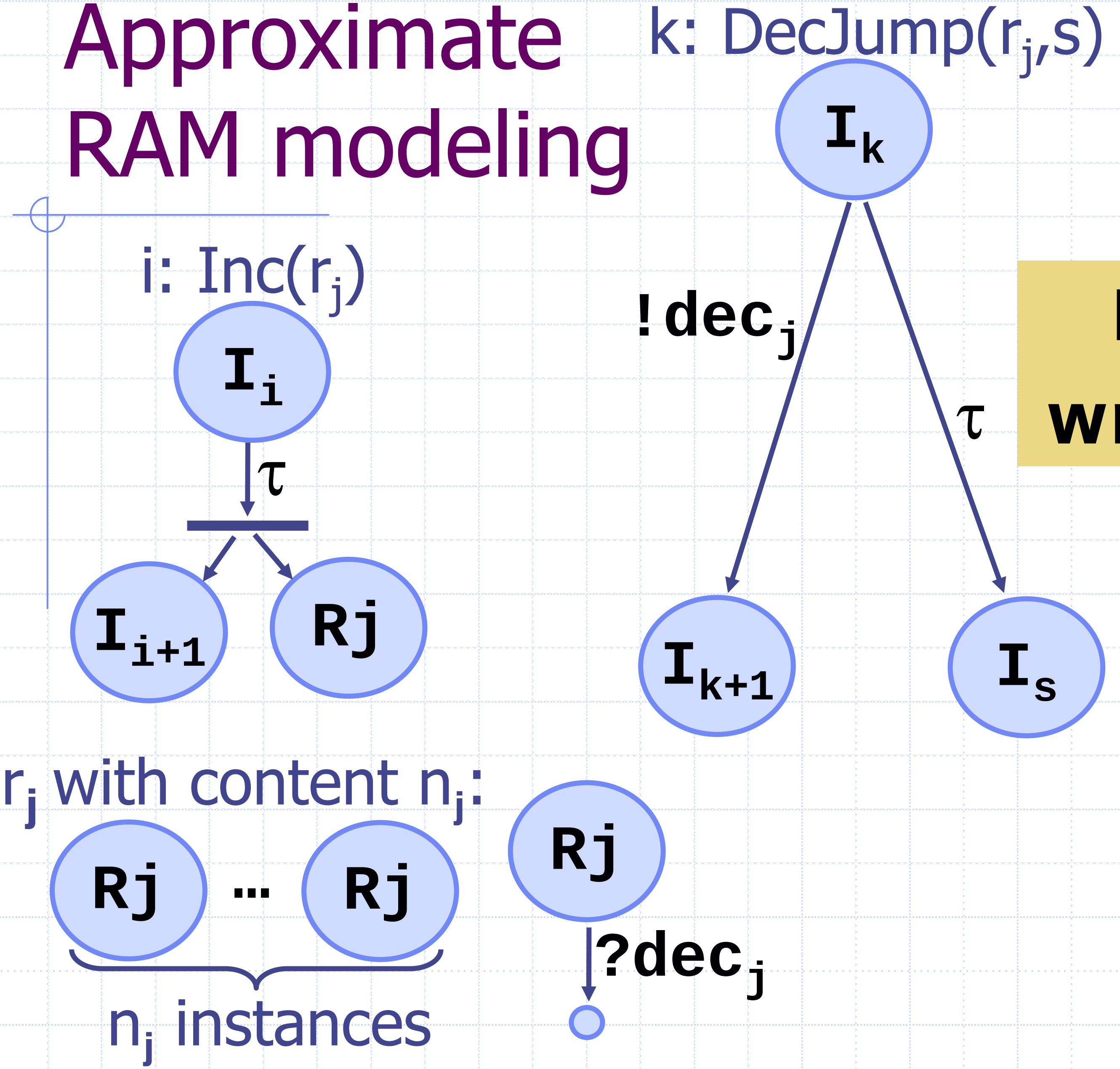
B+C → C+C

C+A → A+A

Is probabilistic universal termination decidable?

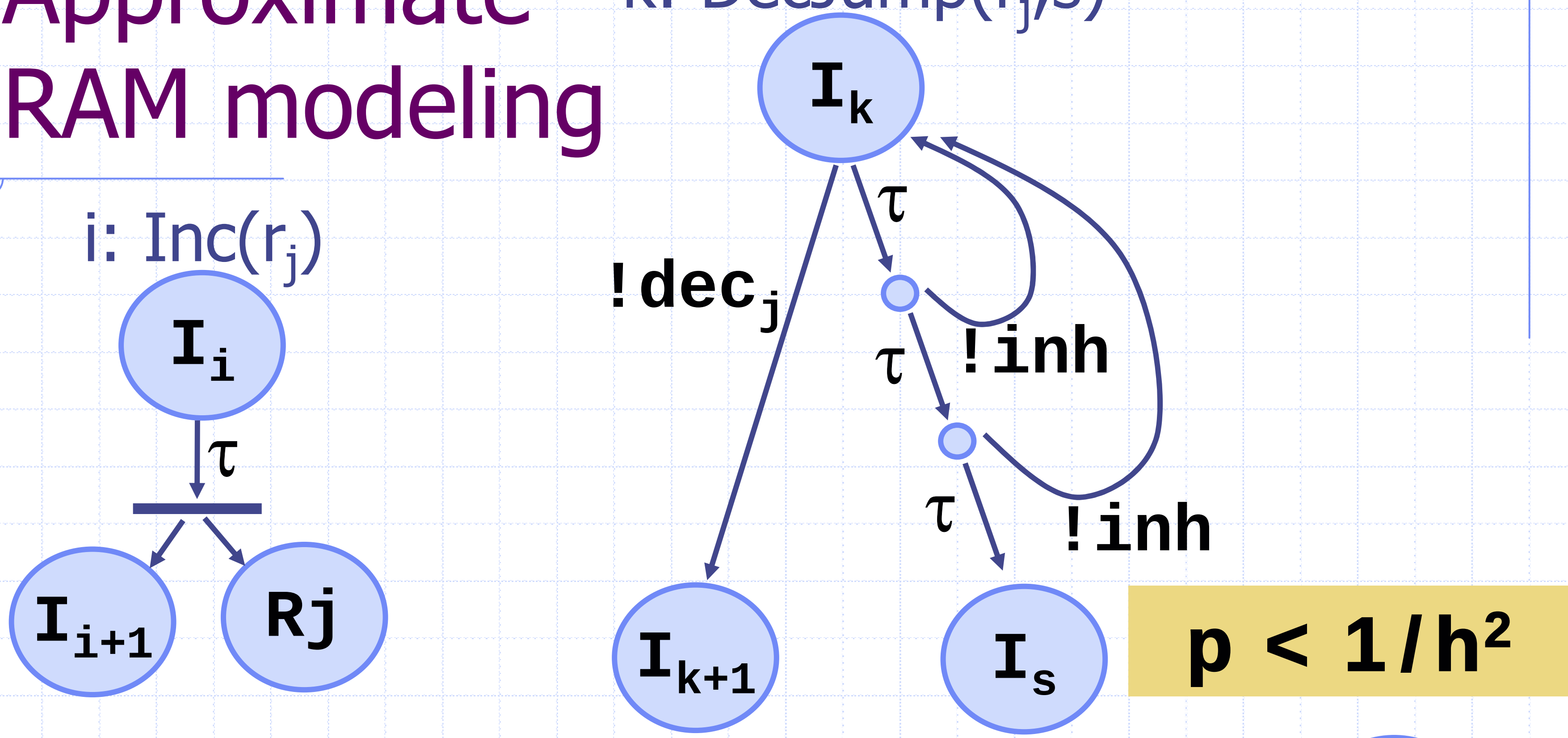
- It is undecidable [Concur08]
- The overall proof includes the proof of the following interesting result:
  - even if RAMs cannot be deterministically modeled in CGF (remember Petri nets modeling of CGF), they can be probabilistically approximated up to any arbitrarily small error ε

Approximate RAM modeling



Problem: wrong jump!

Approximate RAM modeling

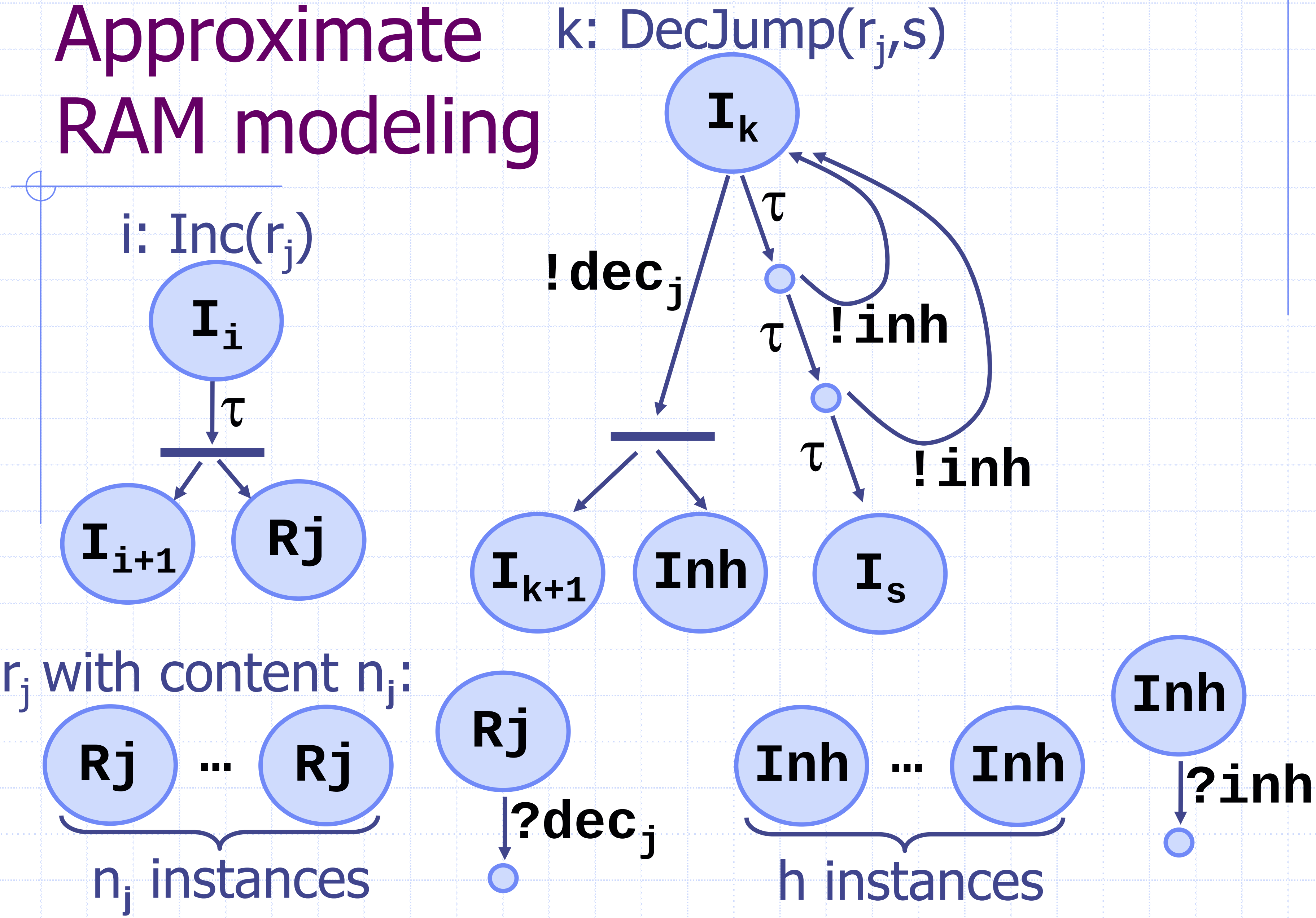


$p < 1/h^2$

r_j with content n :

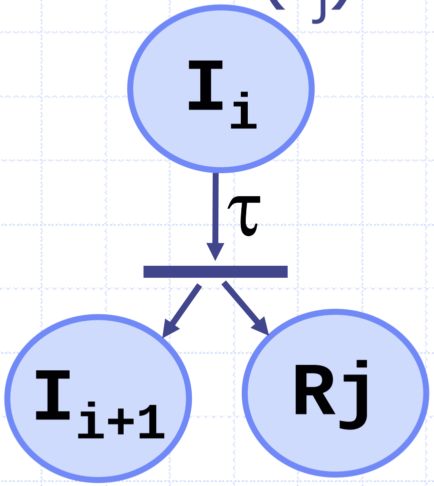
But in an unbounded computation, with infinitely many DecJump's, the prob. of a wrong jump is 1

Approximate RAM modeling

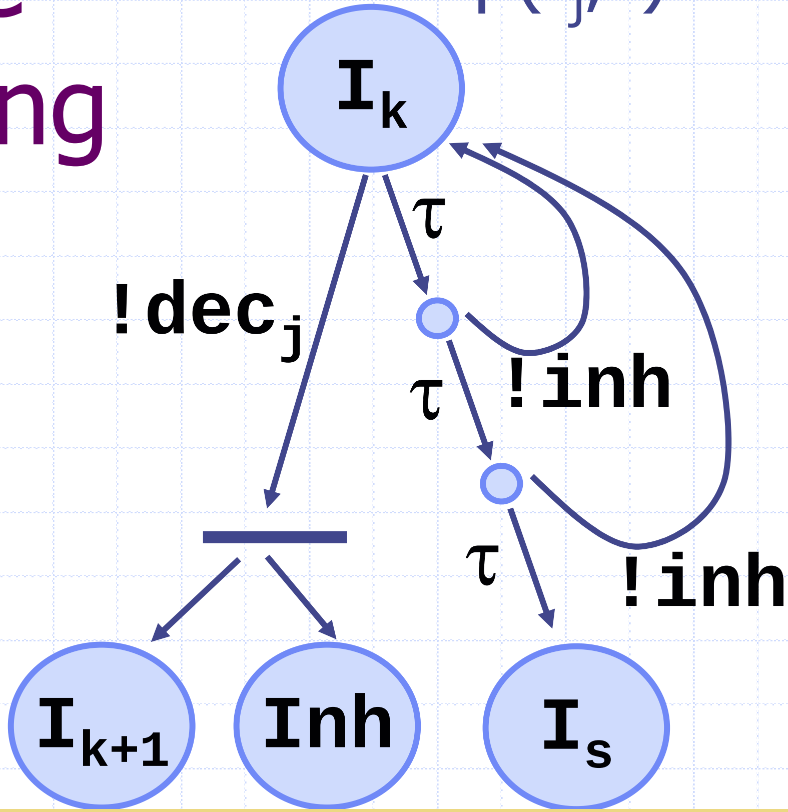


Approximate RAM modeling

i: Inc(r_j)



k: DecJump(r_j,s)



r_j wit

Incrementing the occurrences of Inh the prob. of a wrong jump is

$$< \sum_{k=h}^{\infty} \frac{1}{k^2}$$

Related work

- Magnasco. ***Chemical Kinetics is Turing Universal***. *Phys Rev Lett*. 1997
  - Exploit different reaction rates to model "finite logical circuits with unbounded memory" using unbounded chemical species
- Liekens and Fernando. ***Turing Complete Catalytic Particle Computers***. In Proc. *ECAL'07*. 2007
  - Approximate bounded computations of RAMs
- Soloveichik et al. ***Computation with Finite Stochastic Chemical Reaction Networks***. In *Nat. Computing*. 2008
  - Approximate also unbounded computations of RAMs

References

- Cardelli. **On process rate semantics.** To appear in *Theoretical Computer Science*. 2008
  - Definition of CGF and proof of equivalence with chemical kinetics
- Cardelli. **Artificial Biochemistry.** In Proc. *Algorithmic Bioprocesses '08*. To appear in LNCS. 2008
  - Informal introduction of association/dissociation mechanisms
- Cardelli and Zavattaro. **On the computational power of biochemistry**. In Proc. *AB'08*. To appear in LNCS. 2008
  - Definition of BGF and proof of Turing completeness
- Zavattaro and Cardelli. **Termination problems in chemical kinetics.** In Proc. *Concur'08*. To appear in LNCS. 2008
  - Decidability and nondecidability of nondeterministic and probabilistic versions of properties in CGF