On the Computational Strength of Calculi for (Bio)Chemistry

Based on joint works with:

L. Cardelli (Microsoft Research Cambridge – UK)
A. Credi, M. Garavelli, S. Silvi (Chemistry dep. Univ. of Bologna)
C. Di Giusto, M. Gabbrielli, C. Laneve (Univ. of Bologna)
G. Delzanno (Univ. of Genova)
S. Pradalier (Ecole Polytechnique, Paris)

done within the UniBo Strategic Project CompReNDe
(Compositional and executable Representations of Nano Devices)

Gianluigi Zavattaro
University of Bologna
zavattar@cs.unibo.it
http://cs.unibo.it/~zavattar

On the Computational Strength of Calculi for (Bio)Chemistry

$A \rightarrow^{r} C_1+\ldots+C_n \quad \Longleftrightarrow \quad A ::= \tau@r;C_1|\ldots|C_n + b@s;\mathbf{0}$

$A+B \rightarrow^{s} D_1+\ldots+D_m \quad \Longleftrightarrow \quad B ::= \overline{b}@s;D_1|\ldots|D_m$

What is the computational power of (Bio)Chemistry?

Gianluigi Zavattaro
University of Bologna
zavattar@cs.unibo.it
http://cs.unibo.it/~zavattar

Plan of the talk

- Chemistry as a Computational Model
  - ... not a new issue
- Chemical Ground Form (CGF) [Cardelli08]
  - ... a new way to analyze chemistry
- Computational Strength of Chemistry [Concur08]
  - Non Turing universal but...
  - Turing complete formalisms can be approximated with any degree of precision
- From Chemistry to...
  - Biochemistry: BGF [AB08]
  - Nano devices: nano-K [CMSB07,TCS08]
- Computational strength of BGF and K-calculi [CMSB09]
- Concluding remarks

Is Chemistry Turing powerful?

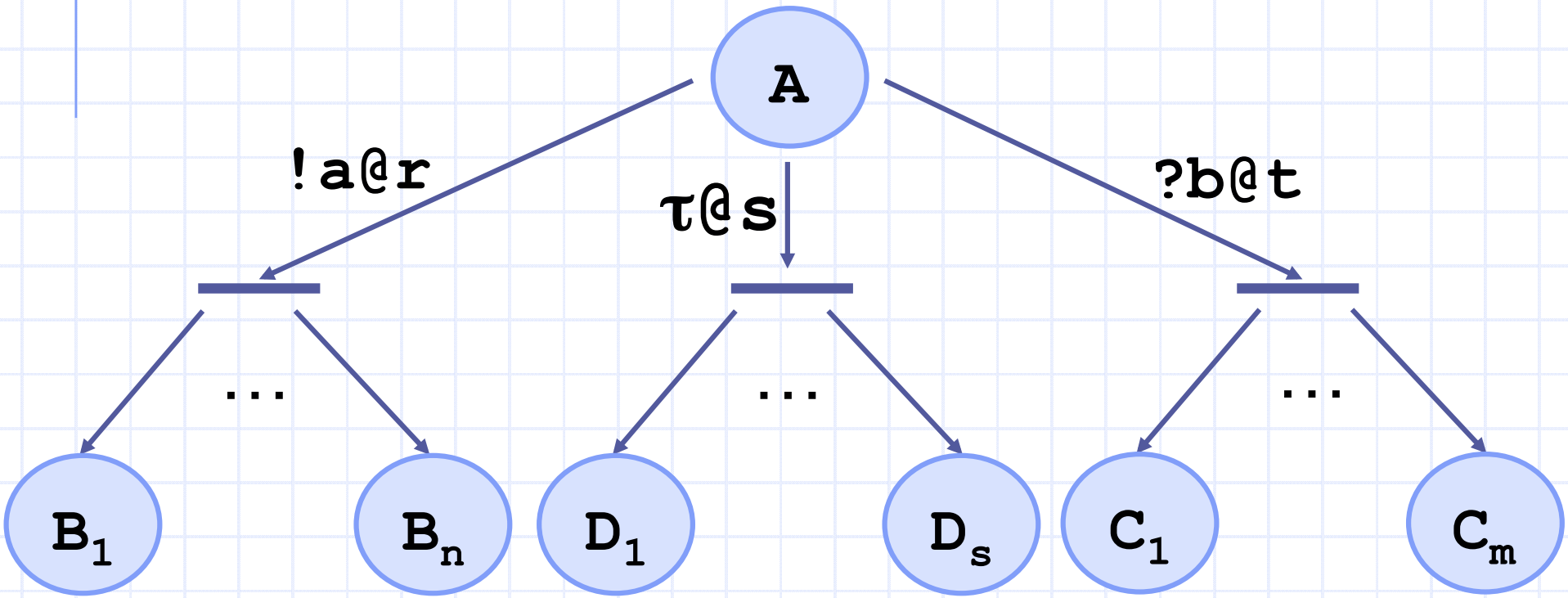
- Magnasco. ***Chemical Kinetics is Turing Universal***. *Phys Rev Lett*. 1997
  - Answer: **YES**… but justification not convincing (only Digital Computers with bounded memory are considered)
- Liekens and Fernando. ***Turing Complete Catalytic Particle Computers***. In Proc. *ECAL'07*. 2007
  - Answer: **YES**… but justification not convincing (only Minsky Machines with bounded computation are considered)
- Soloveichik et al. ***Computation with Finite Stochastic Chemical Reaction Networks***. *Nat. Computing*. 2008
  - Answer: **NO**… but all Minsky and Turing Machines can be at least approximated with any given degree of precision

Plan of the talk

- Chemistry as a Computational Model
  - ... not a new issue
- **Chemical Ground Form (CGF) [Cardelli08]**
  - **... a new way to analyze chemistry**
- Computational Strength of Chemistry [Concur08]
  - Non Turing universal but...
  - Turing complete formalisms can be approximated with any degree of precision
- From Chemistry to...
  - Biochemistry: BGF [AB08]
  - Nano devices: nano-K [CMSB07,TCS08]
- Computational strength of BGF and K-calculi [CMSB09]
- Concluding remarks

Chemical Ground Forms

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

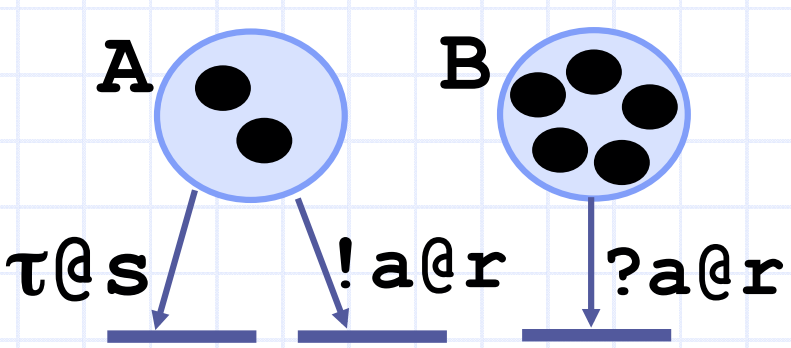


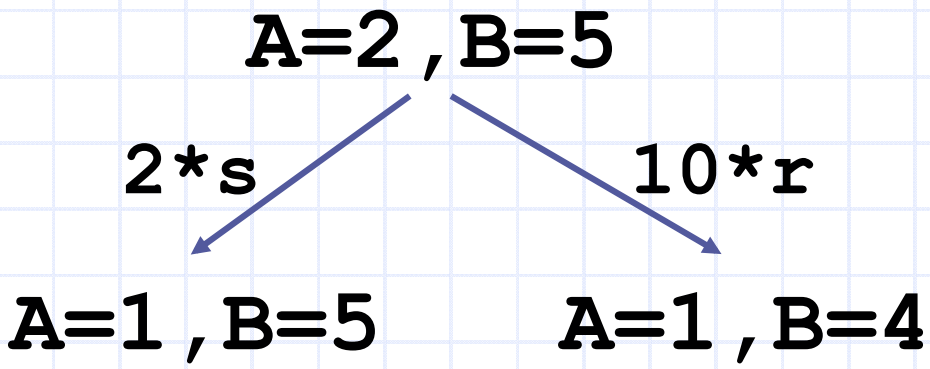
Stochastic semantics

- Actions take (an exponentially distributed amount of) time
  - 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}**

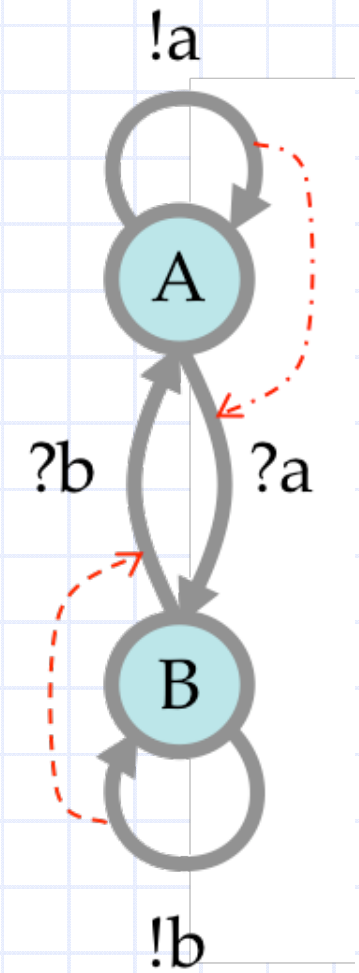
Collective stochastic semantics

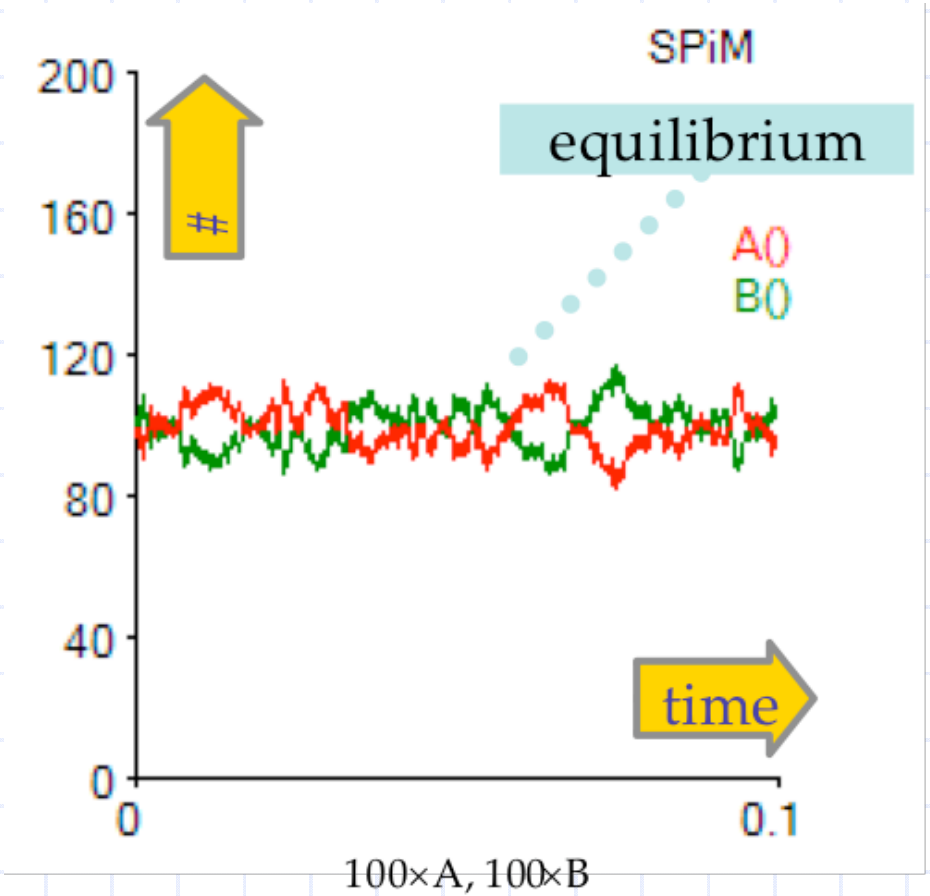
- To faithfully model chemistry, the rates of distinct transitions with the same effect **must be added**



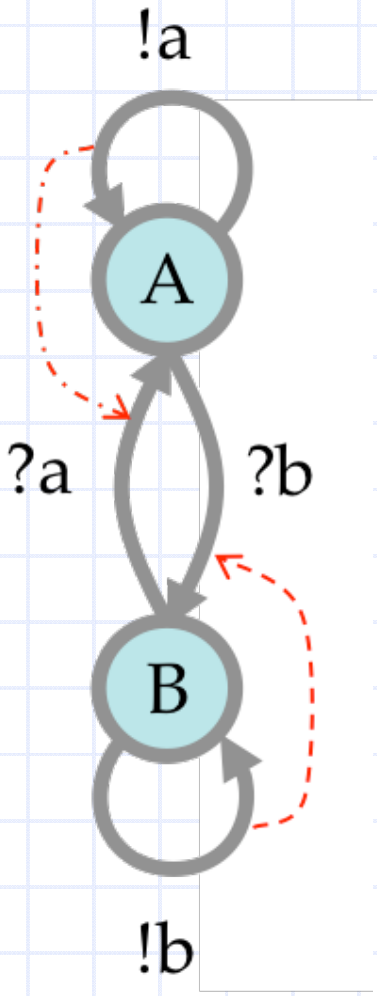


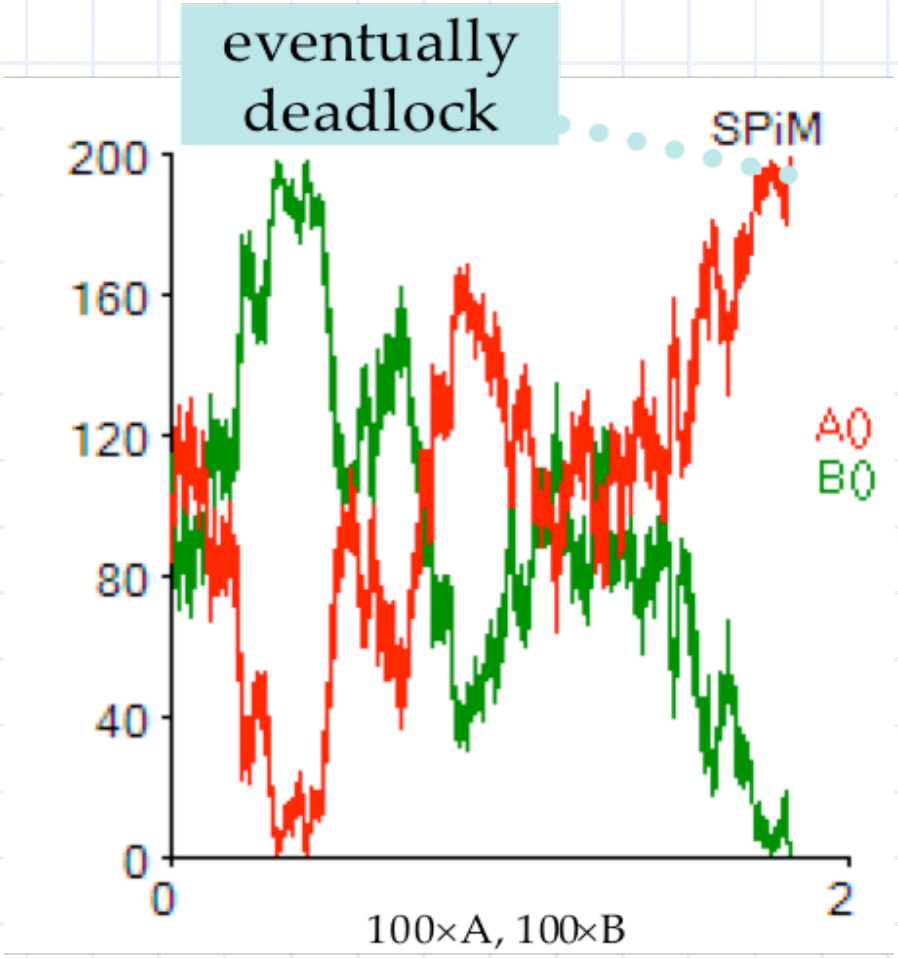
Example: Celebrities…



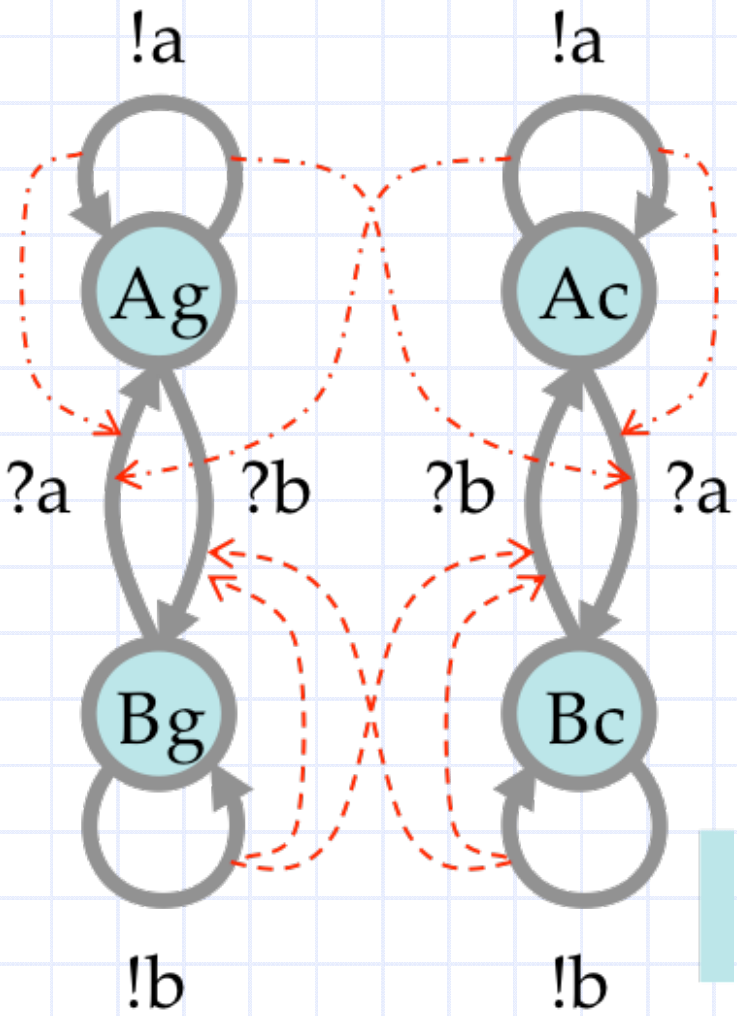


Example: Groupies...



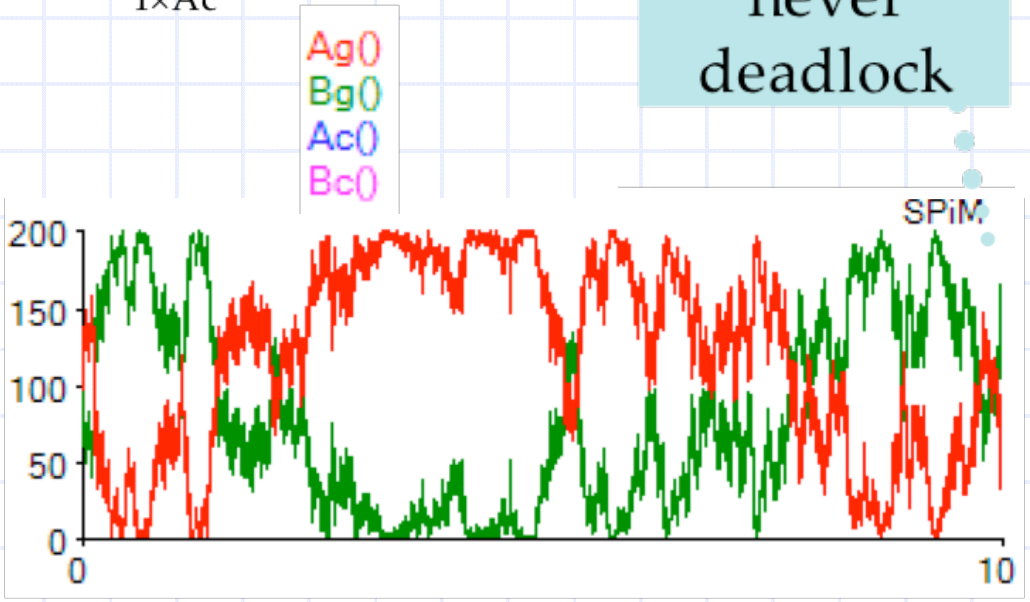


Example:
Groupies with one Celebrity…



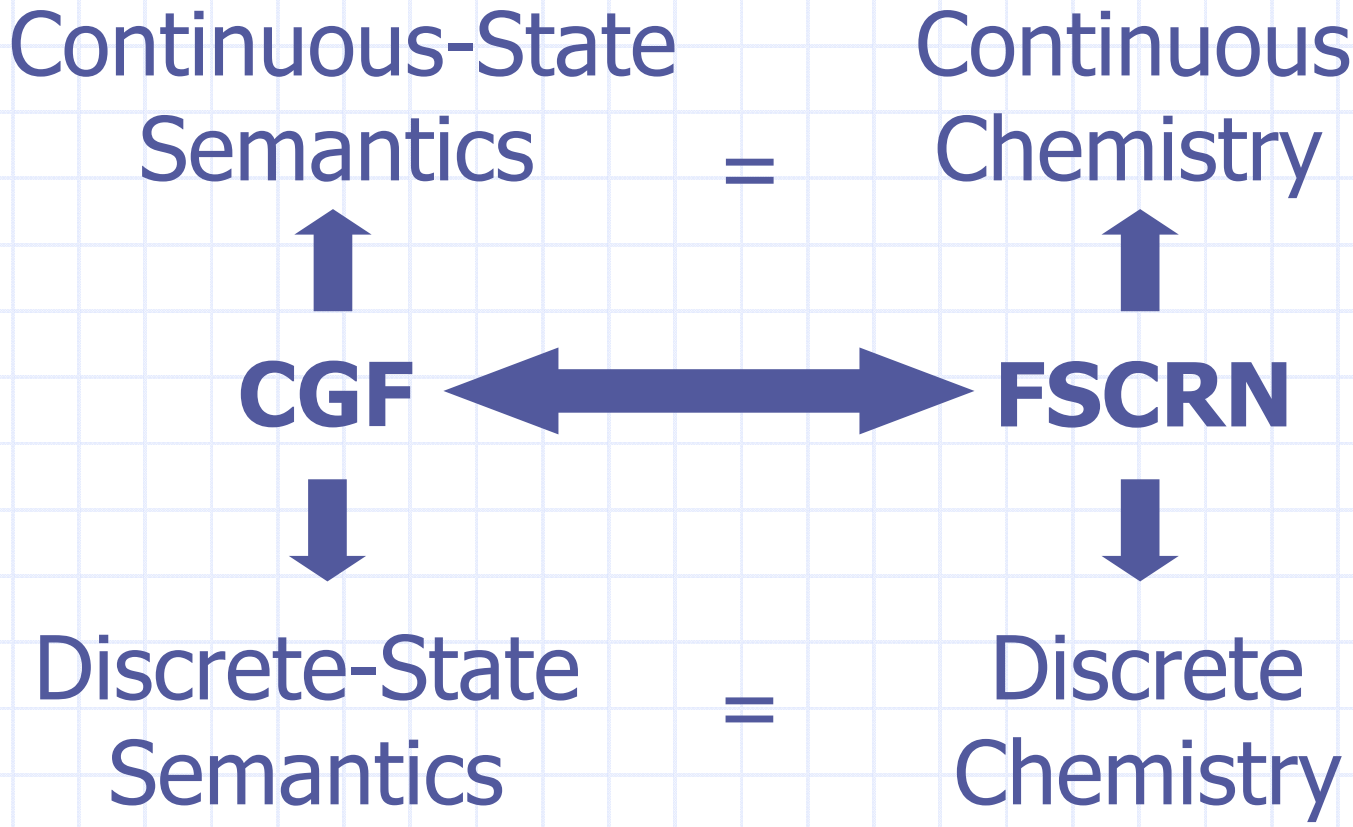
100×Ag, 100×Bg
1×Ac

never deadlock



CGF = FSCRN

[Cardelli08]



Plan of the talk

- Chemistry as a Computational Model
  - ... not a new issue
- Chemical Ground Form (CGF) [Cardelli08]
  - ... a new way to analyze chemistry
- **Computational Strength of Chemistry** [Concur08]
  - **Non Turing universal but...**
  - Turing complete formalisms can be approximated with any degree of precision
- From Chemistry to...
  - Biochemistry: BGF [AB08]
  - Nano devices: nano-K [CMSB07,TCS08]
- Computational strength of BGF and K-calculi [CMSB09]
- Concluding remarks

Non Turing universality

- The (existential version of) the Halting problem is decidable in CGF
- **Proof**:
  - We first define an encoding of CGF into (nondeterministic) **Petri nets**
  - Theorem:
    Given a CGF process P, it can reach a halted configuration (with prob. > 0) **iff** the net that encodes P has a deadlock
  - **Deadlock is decidable** in Petri nets

Plan of the talk

- Chemistry as a Computational Model
  - ... not a new issue
- Chemical Ground Form (CGF) [Cardelli08]
  - ... a new way to analyze chemistry
- **Computational Strength of Chemistry** [Concur08]
  - Non Turing universal but...
  - **Turing complete formalisms can be approximated with any degree of precision**
- From Chemistry to...
  - Biochemistry: BGF [AB08]
  - Nano devices: nano-K [CMSB07,TCS08]
- Computational strength of BGF and K-calculi [CMSB09]
- Concluding remarks

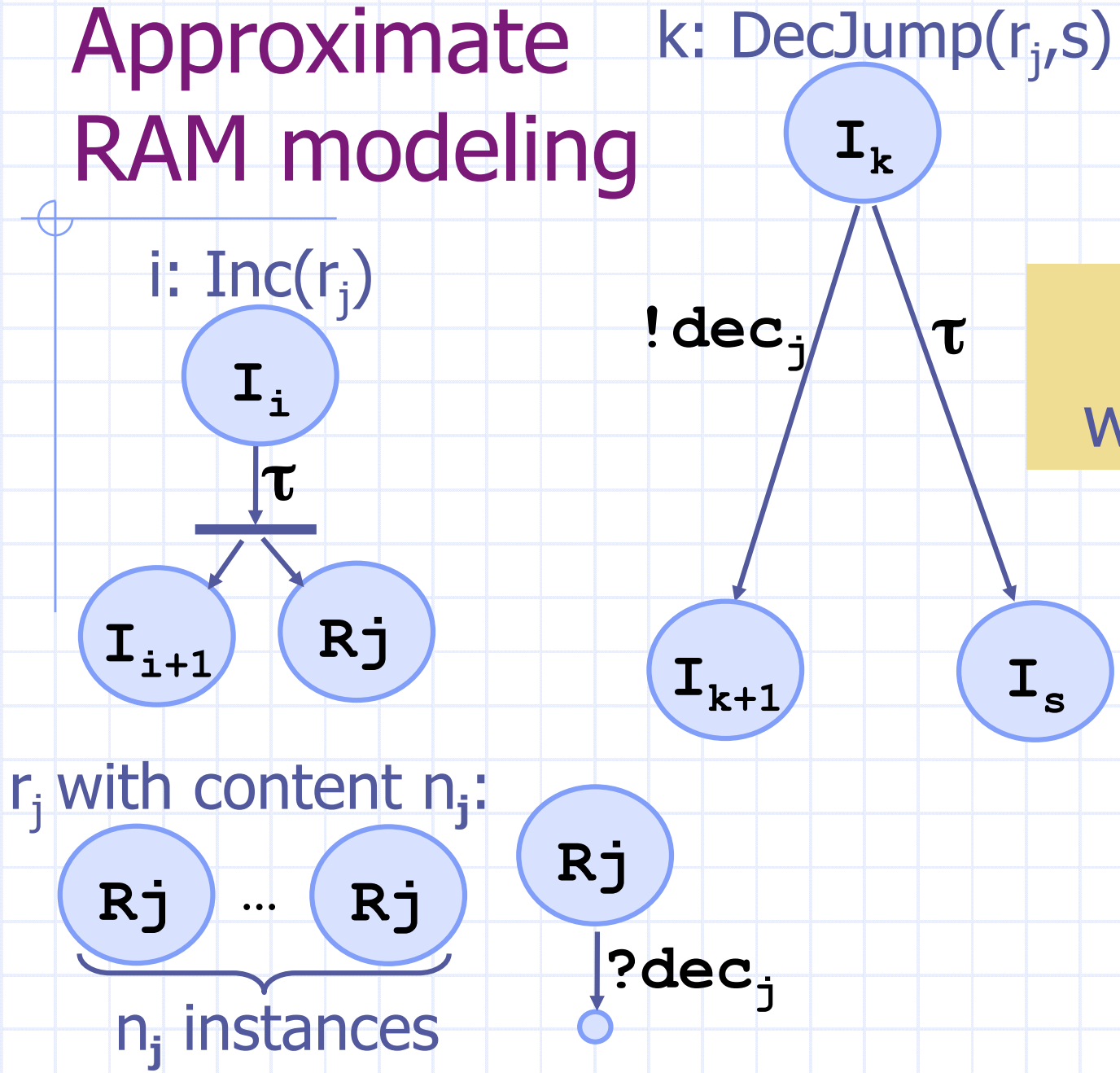
Approximation of Turing complete formalisms in CGF

- We approximate (for any degree of precision) **Random Access Machines** (RAMs)
- RAMs [Min67]:
  - **Registers**: r_1 ... r_n hold natural numbers
  - **Program**: sequence of indexed 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

RAM encoding

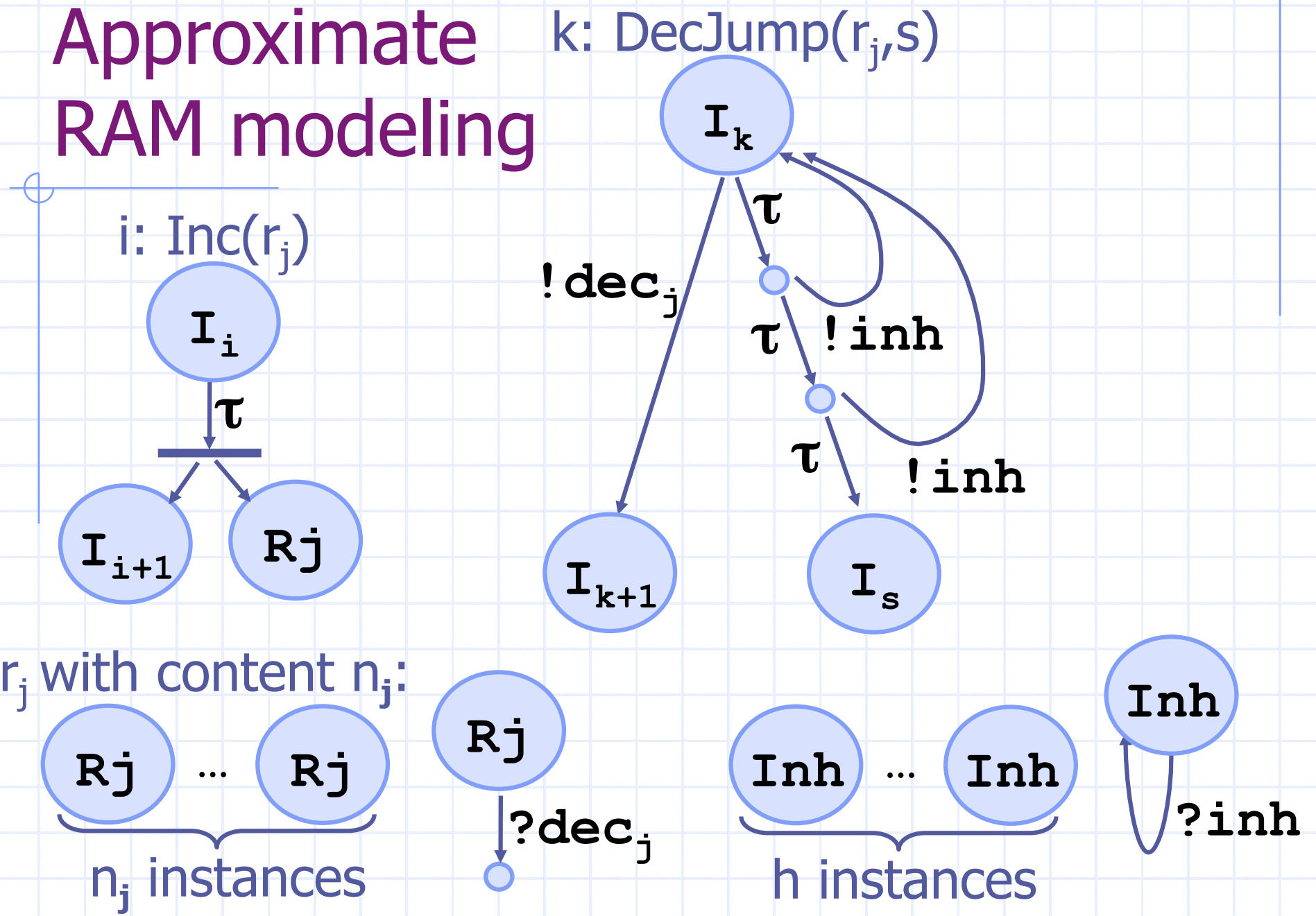
- RAMs **cannot** be faithfully modeled by a CGF process
  - otherwise (by decidability of existential term. in CGF) RAM termination is decidable
- **Main theorem:**
  Given $\varepsilon > 0$ and a RAM, it **can** be modeled by a CGF process that includes also wrong computations, but the probability a wrong computation is scheduled is smaller than ε

Approximate RAM modeling

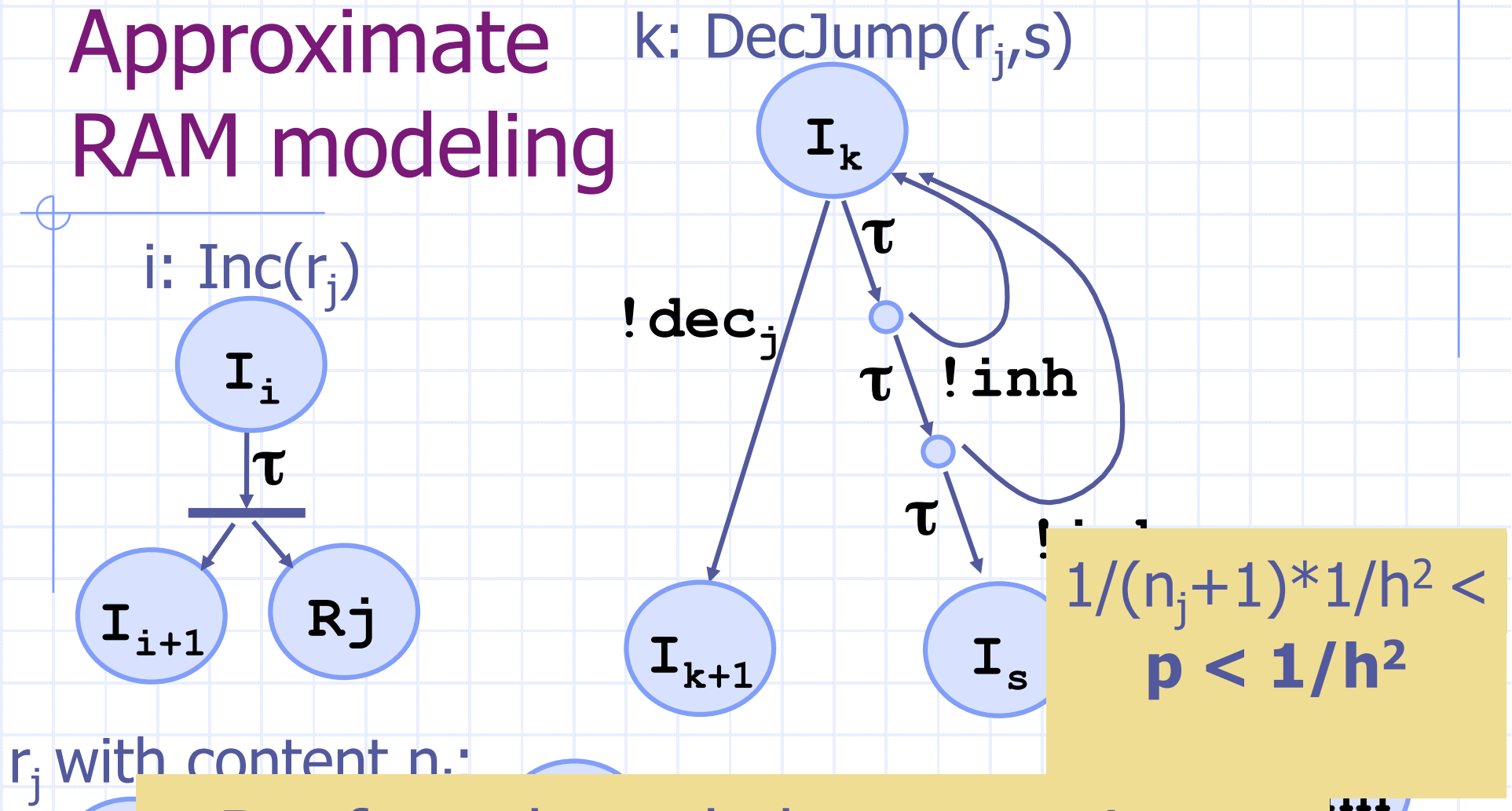


Problem:
wrong jump!

Approximate RAM modeling



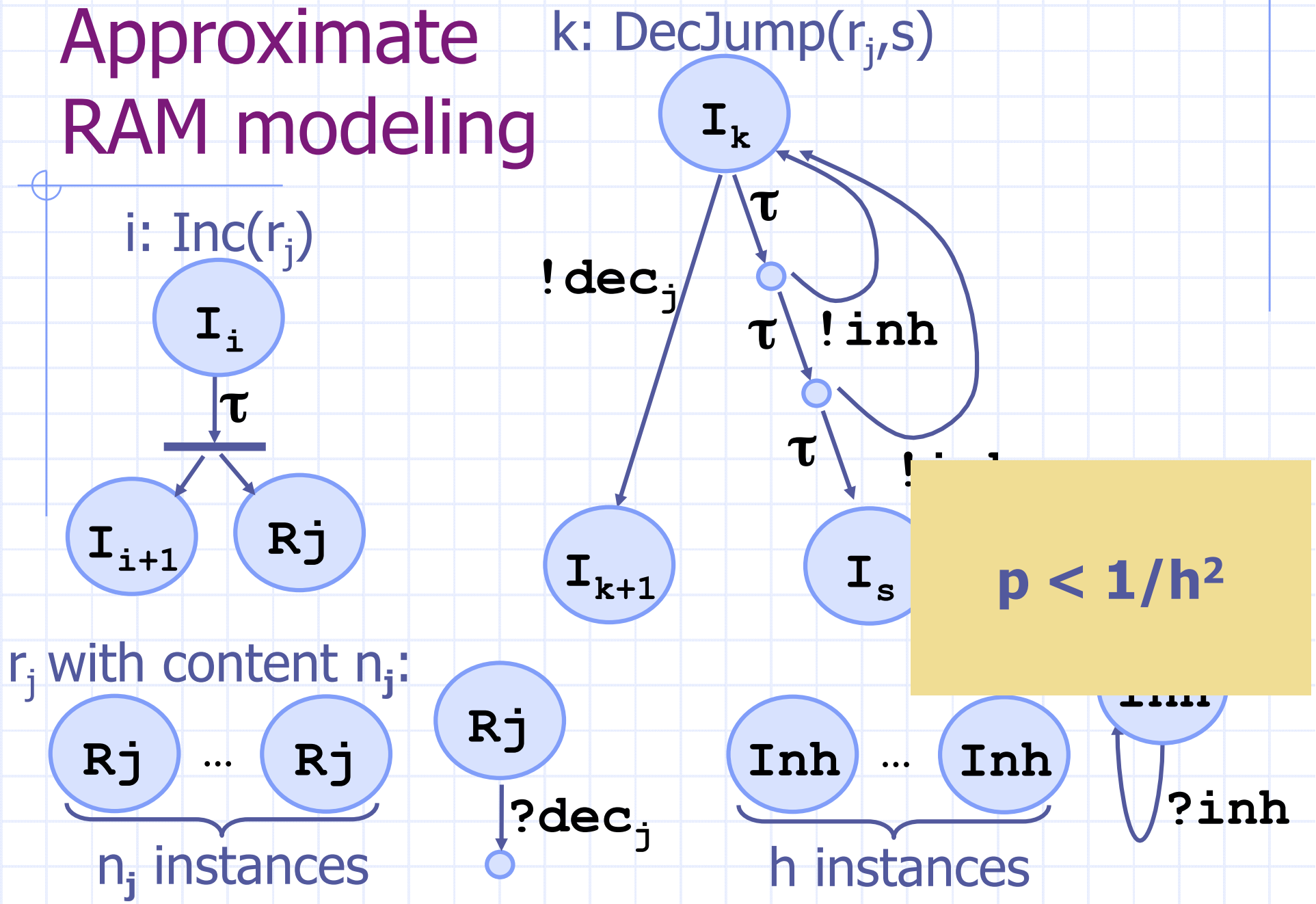
Approximate RAM modeling



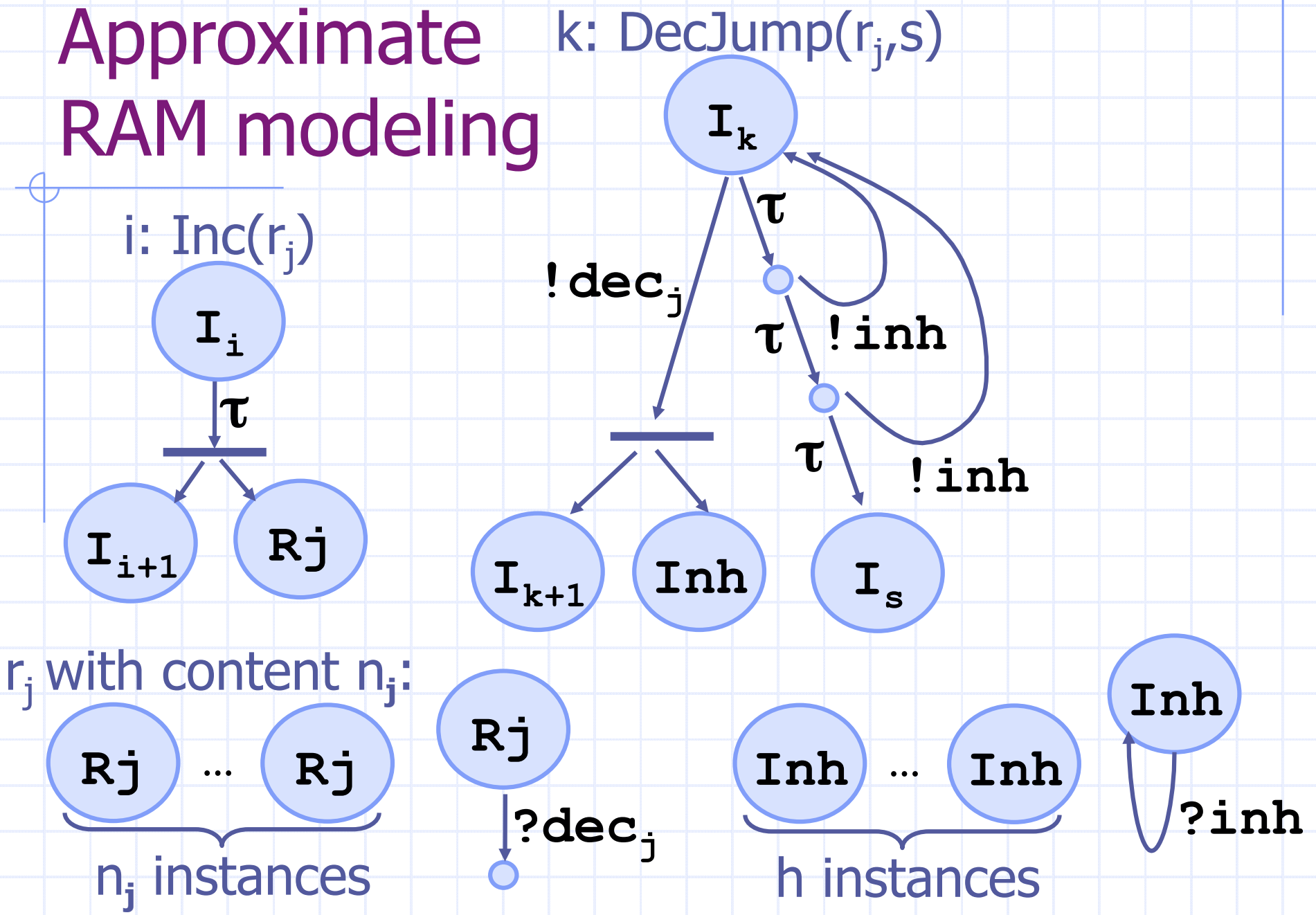
$1/(n_j+1)*1/h^2 < p < 1/h^2$

But for unbounded computations, with infinitely many decrements, the prob. of a wrong jump is 1

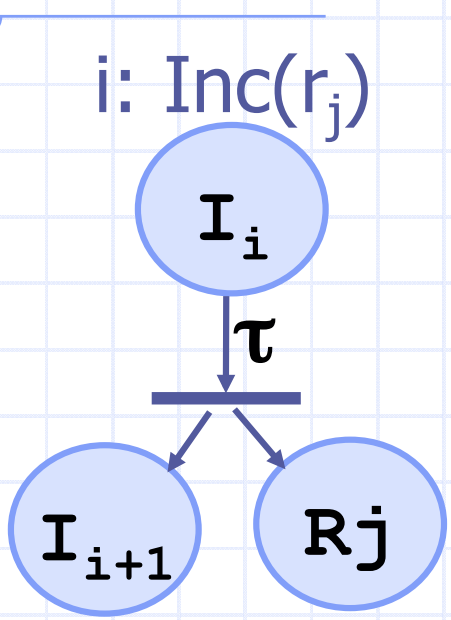
Approximate RAM modeling

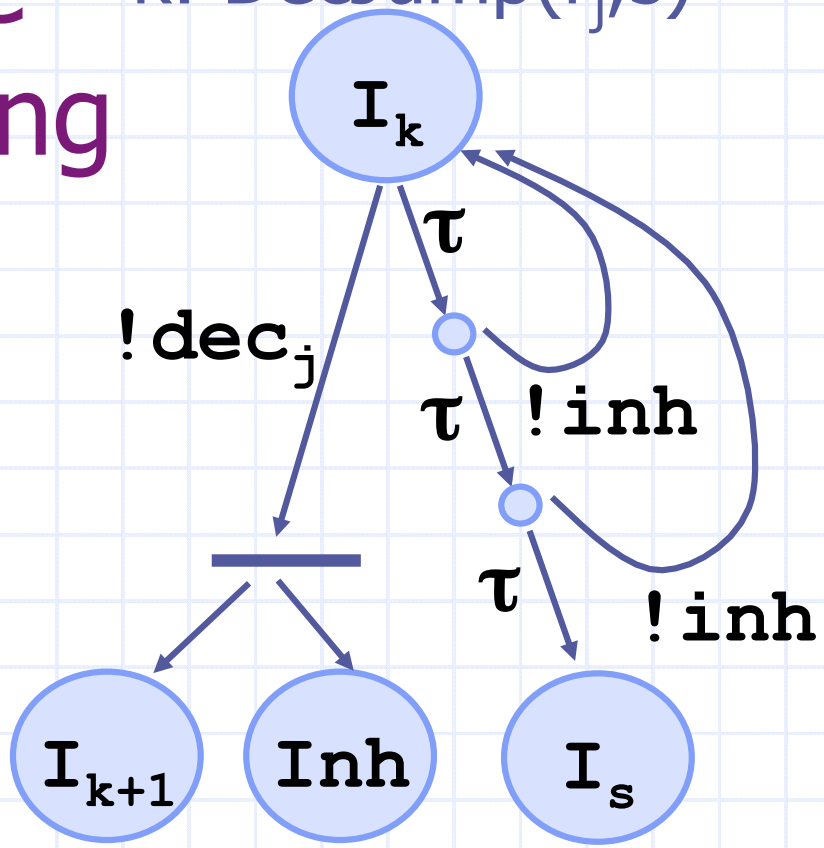


Approximate RAM modeling



Approximate RAM modeling





r_j wi

nh

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

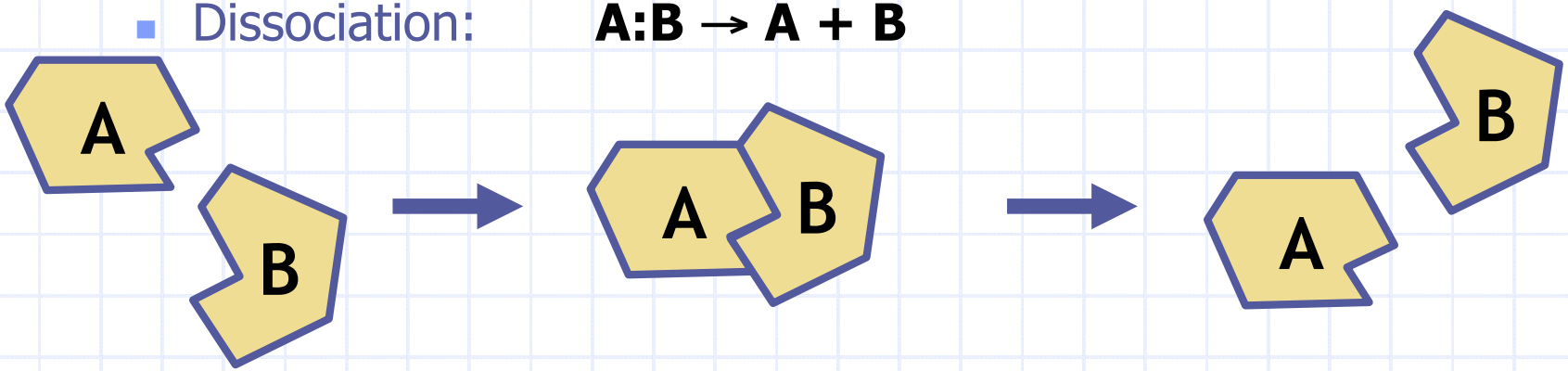
$$p < \Sigma_{i=inith...\infty} 1/i^2$$

Plan of the talk

- Chemistry as a Computational Model
  - ... not a new issue
- Chemical Ground Form (CGF) [Cardelli08]
  - ... a new way to analyze chemistry
- Computational Strength of Chemistry [Concur08]
  - Non Turing universal but...
  - Turing complete formalisms can be approximated with any degree of precision
- **From Chemistry to...**
  - **Biochemistry: BGF [AB08]**
  - Nano devices: nano-K [CMSB07,TCS08]
- Computational strength of BGF and K-calculi [CMSB09]
- Concluding remarks

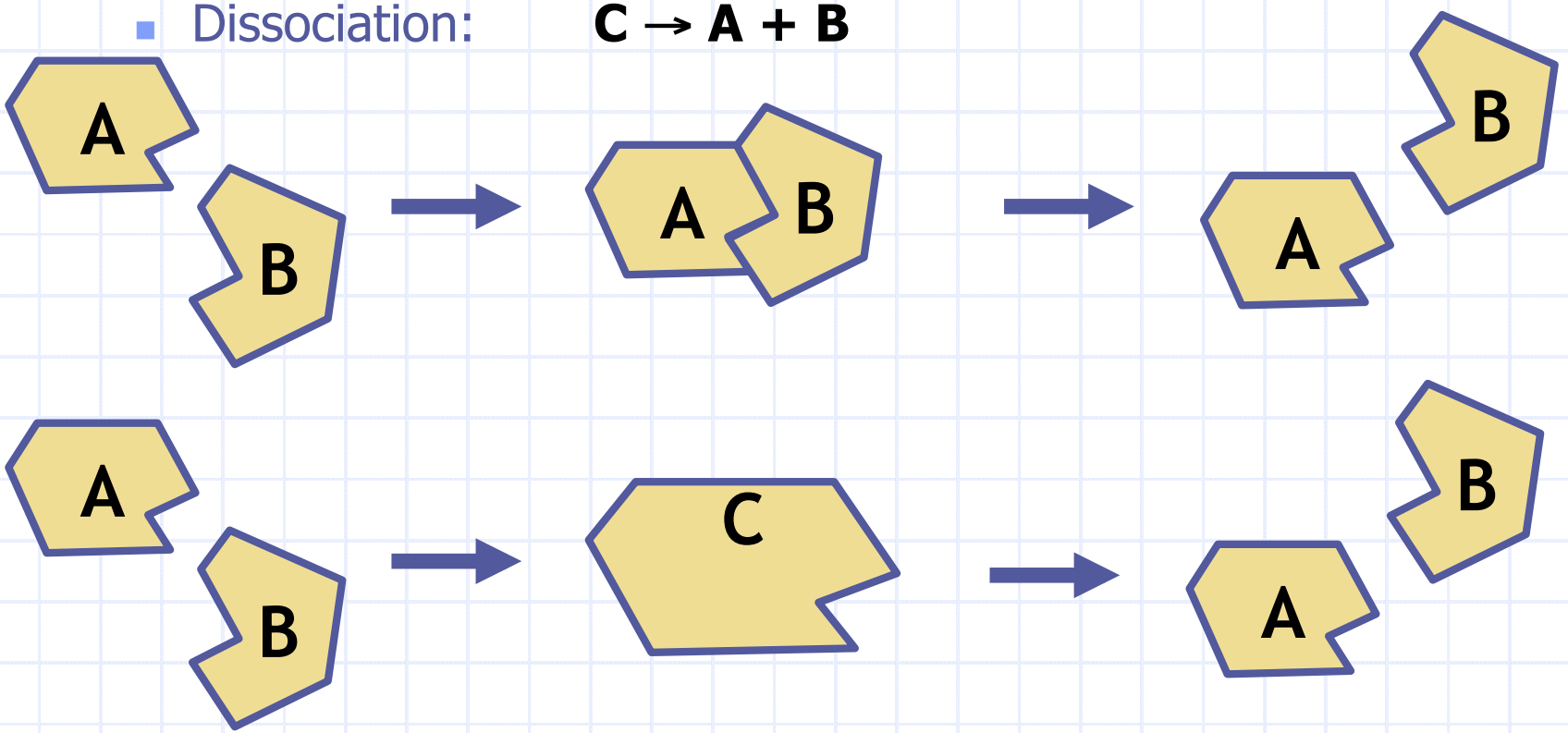
Basic Biochemistry

- Molecules may also form reversible complexes
  - Association: **A + B → A:B**
  - Dissociation: **A:B → A + B**

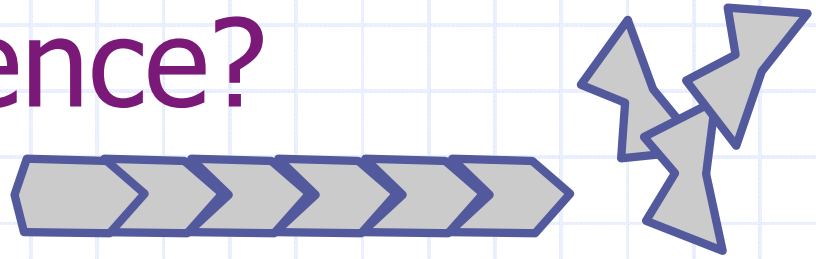


Basic Biochemistry

- Molecules may also form reversible complexes
  - Association: **A + B → C**
  - Dissociation: **C → A + B**



What's the Difference?



Consider linear polymerization:

The "chemical program" for polymerization:

$P_0 + M \rightarrow P_1$

$P_1 + M \rightarrow P_2$

$P_2 + M \rightarrow P_3$

....

an infinite (non-)program

an infinite set of species

$P_{10757} + M \rightarrow P_{10758}$

Such specificity is unreal

But "nature's program" for polymerization has to fit e.g. in the genome, so it cannot be infinite! Clearly, nature must be using a different "language" than basic chemistry:

molecule with convex patch +

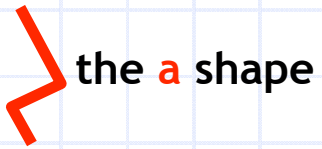
molecule with concave patch →

molecule with convex patch

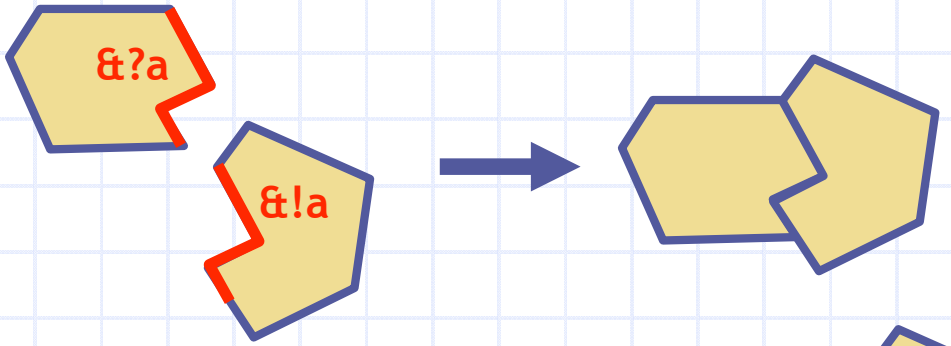
Finite program composed of local rules

Biochemical Ground Form= CGF + Association and Dissociation

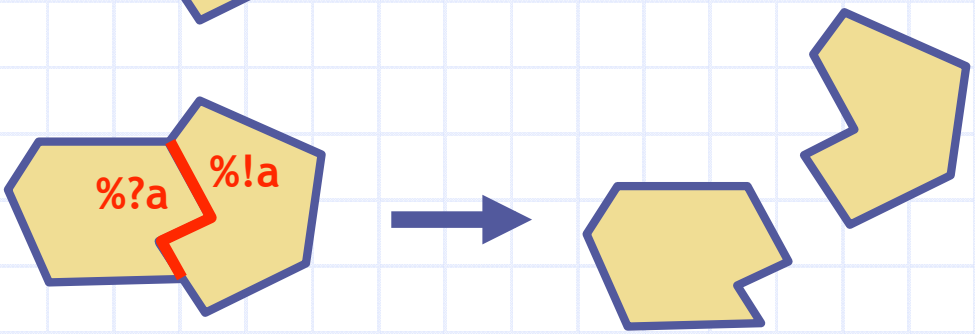
- Association patches are named



- & – association
  - &?a associate
  - &!a co-associate



- % – dissociation
  - %?a dissociate
  - %!a co-dissociate



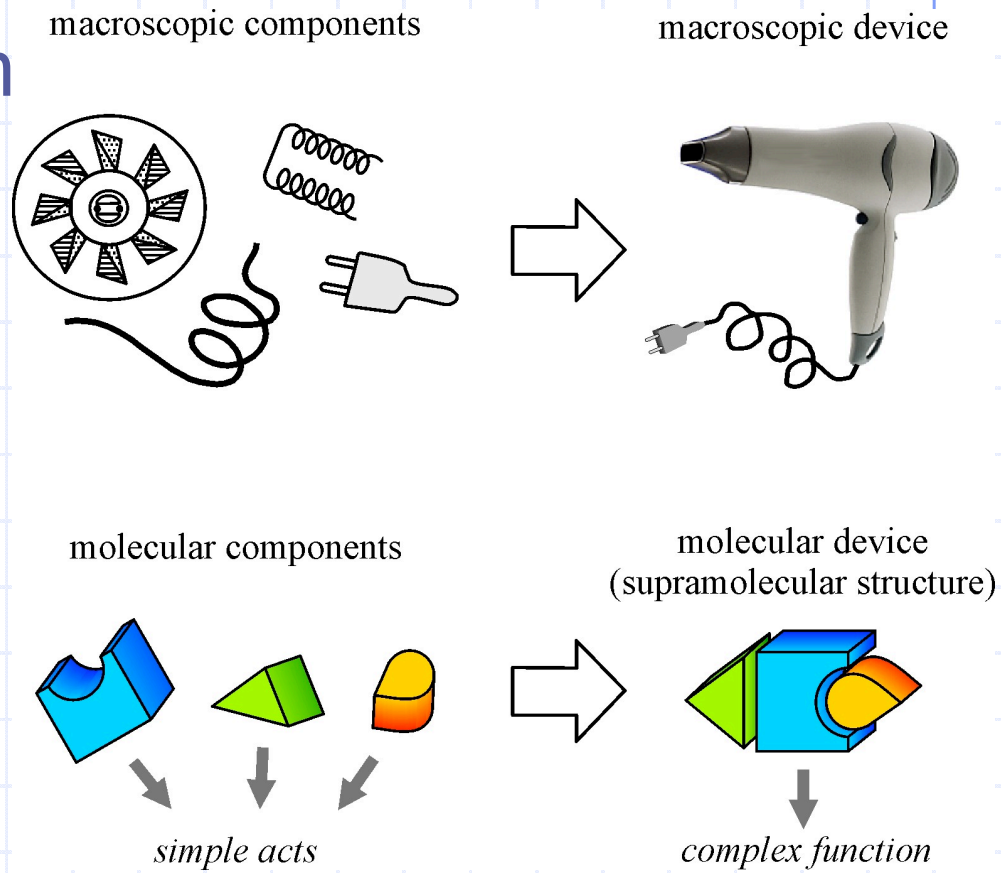
- A given patch can *hold* only one association at a time
- Two molecules can dissociate only if *they* are associated

Plan of the talk

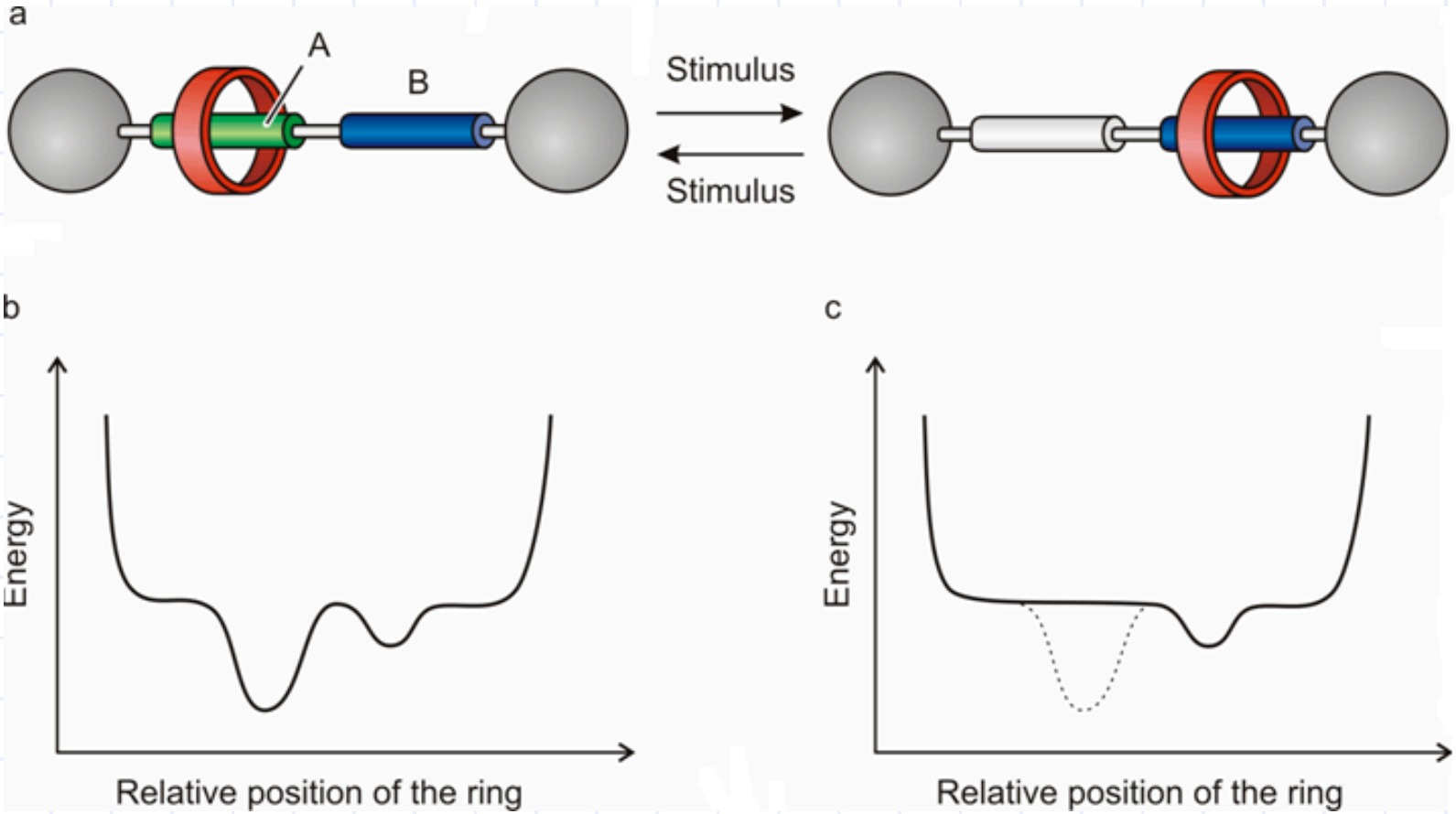
- Chemistry as a Computational Model
  - ... not a new issue
- Chemical Ground Form (CGF) [Cardelli08]
  - ... a new way to analyze chemistry
- Computational Strength of Chemistry [Concur08]
  - Non Turing universal but...
  - Turing complete formalisms can be approximated with any degree of precision
- **From Chemistry to...**
  - Biochemistry: BGF [AB08]
  - **Nano devices: nano-K [CMSB07,TCS08]**
- Computational strength of BGF and K-calculi [CMSB09]
- Concluding remarks

Nano Devices

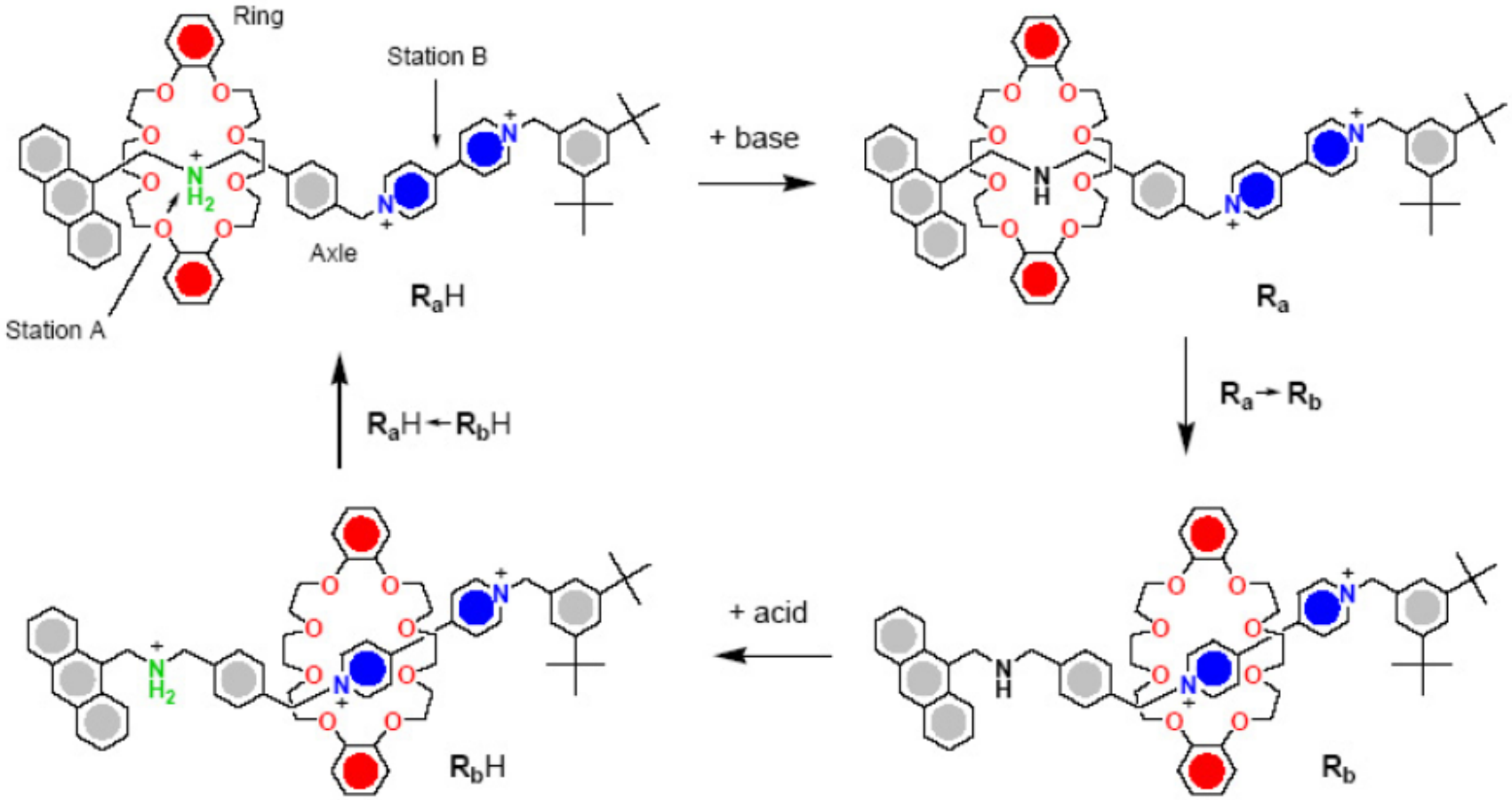
- Supramolecular system
  - Built **bottom-up**: assembling molecular components
  - Used to "implement" basic **functionalities**: binary memories, basic logic gates, mechanical movements, …



Example: [2]rotaxane

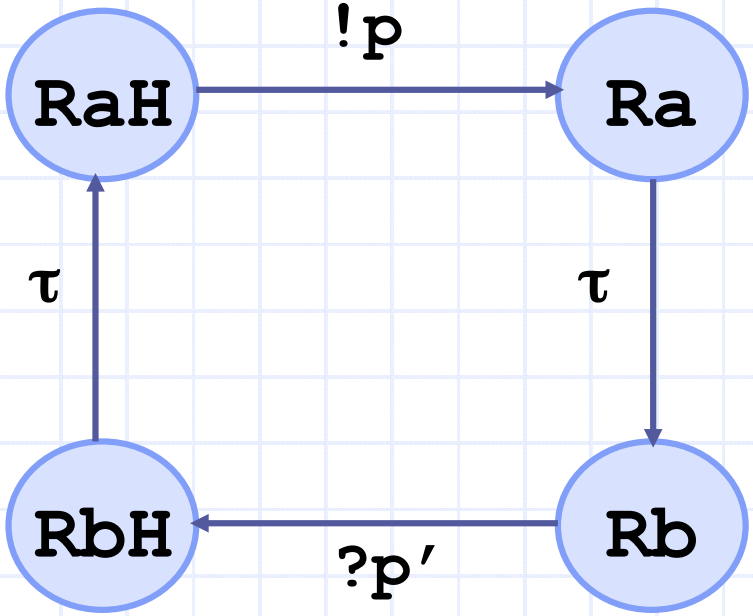


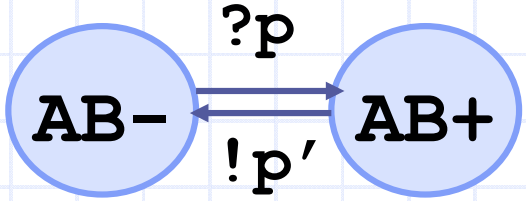
Rotaxane RaH



Modeling RaH in CGF

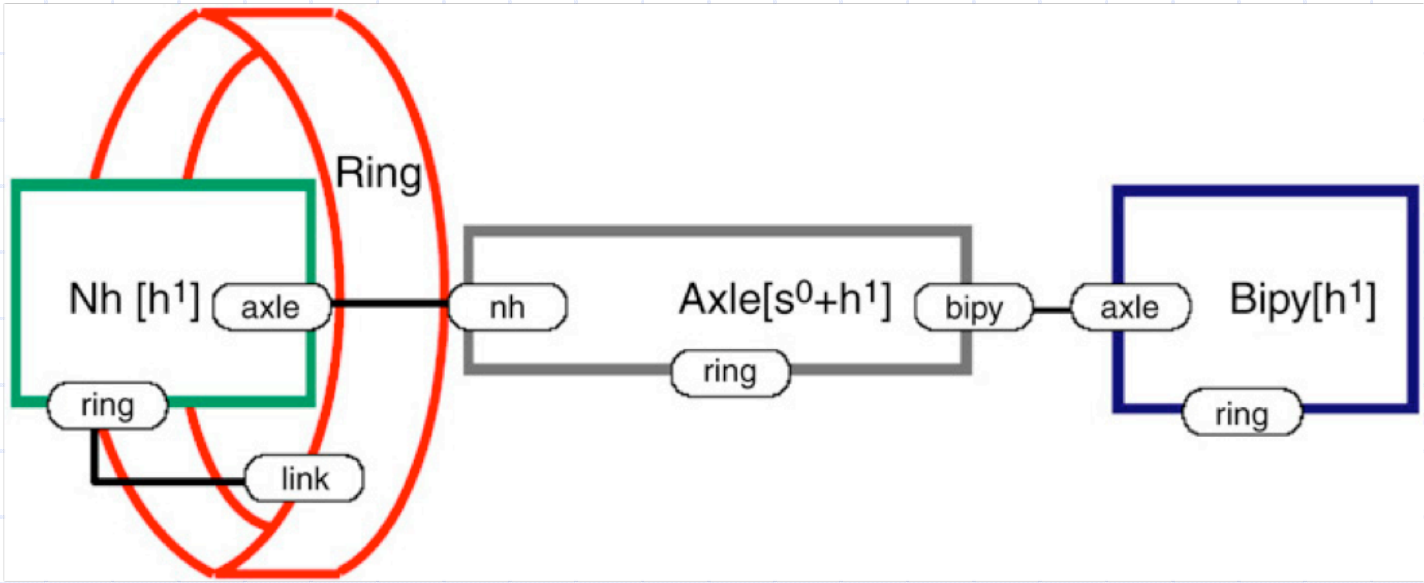
- The RaH process calculus definition





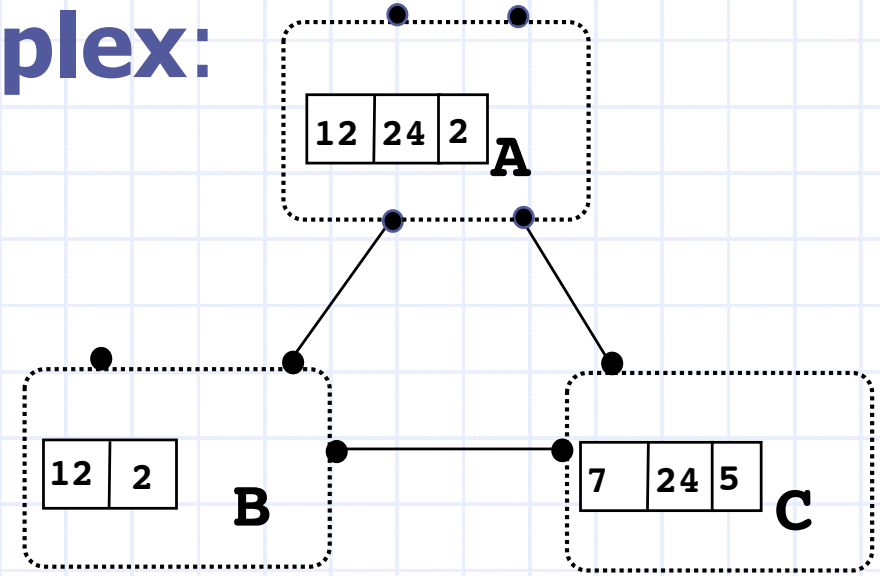
Towards a compositional representation

- In this representation there is no indication of the exploited **components**!
- **Solution**: follow the approach of K-calculi



K-calculi: the nodes

- K is a **graph-rewriting** system where nodes
  - carry **state information** (about phosphorilation, shape,...)
  - have **sites** through which they bind other nodes
- A **complex**:



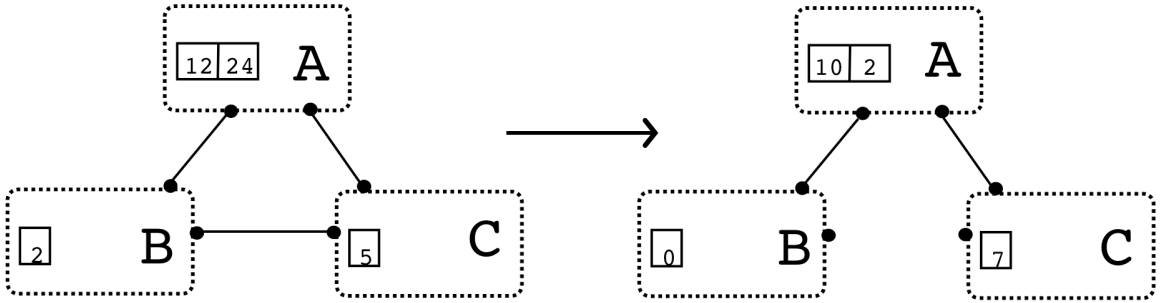
and its syntactic description:

K-calculi: the reactions

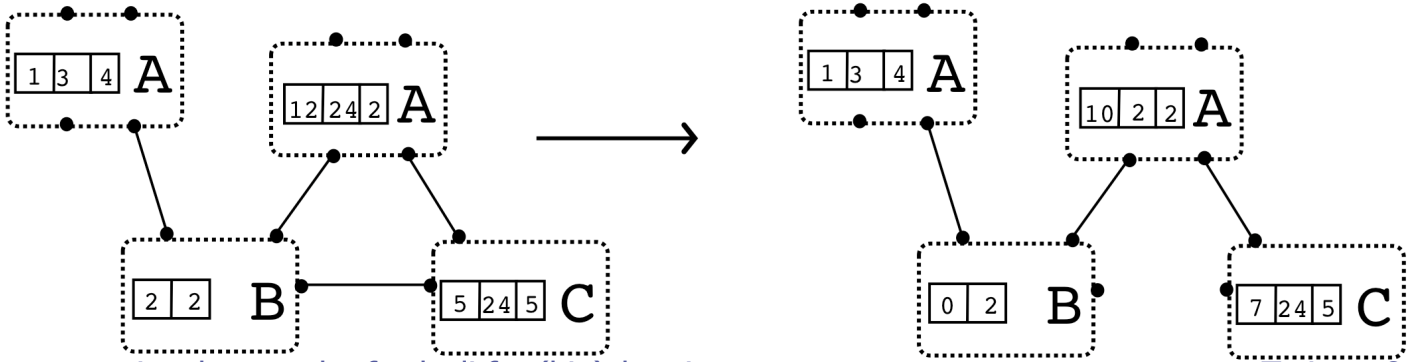
- A reaction rule:

$$A[1^{12}+2^{24}](1^{x}+2^{y}),\ B[1^{2}](2^{x}+3^{z}),\ C[1^{5}](1^{z}+2^{y})$$

$$\rightarrow A[1^{10}+2^{2}](1^{x}+2^{y}),\ B[1^{0}](2^{x}+3),\ C[1^{7}](1+2^{y})$$

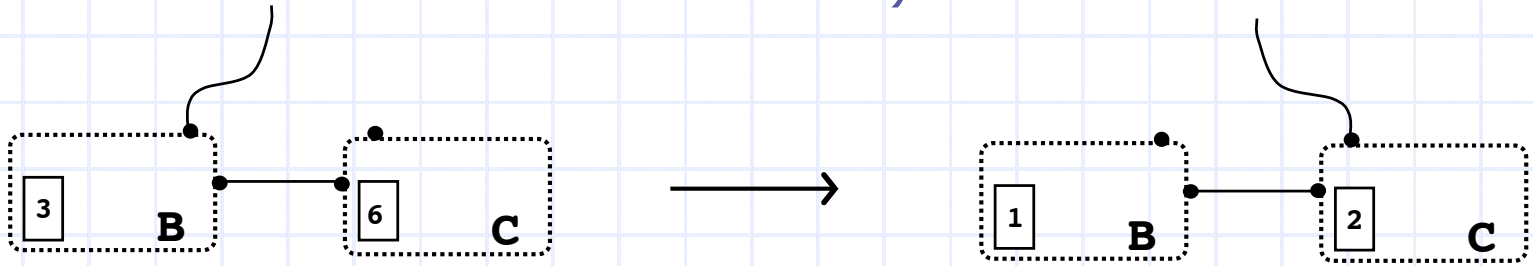


and a possible instantiation:

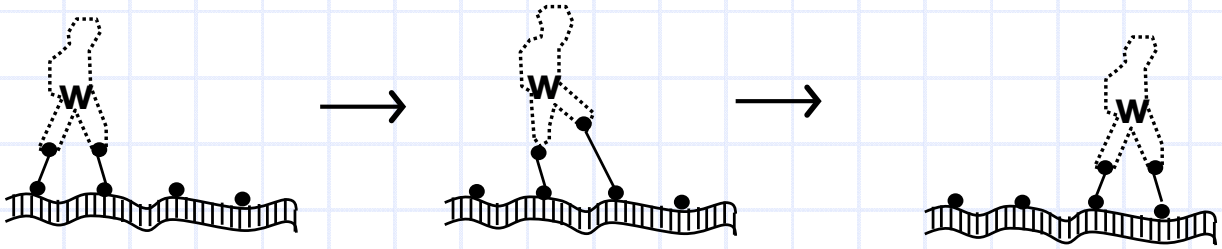


Three types of reactions: creations, destructions and bond flipping

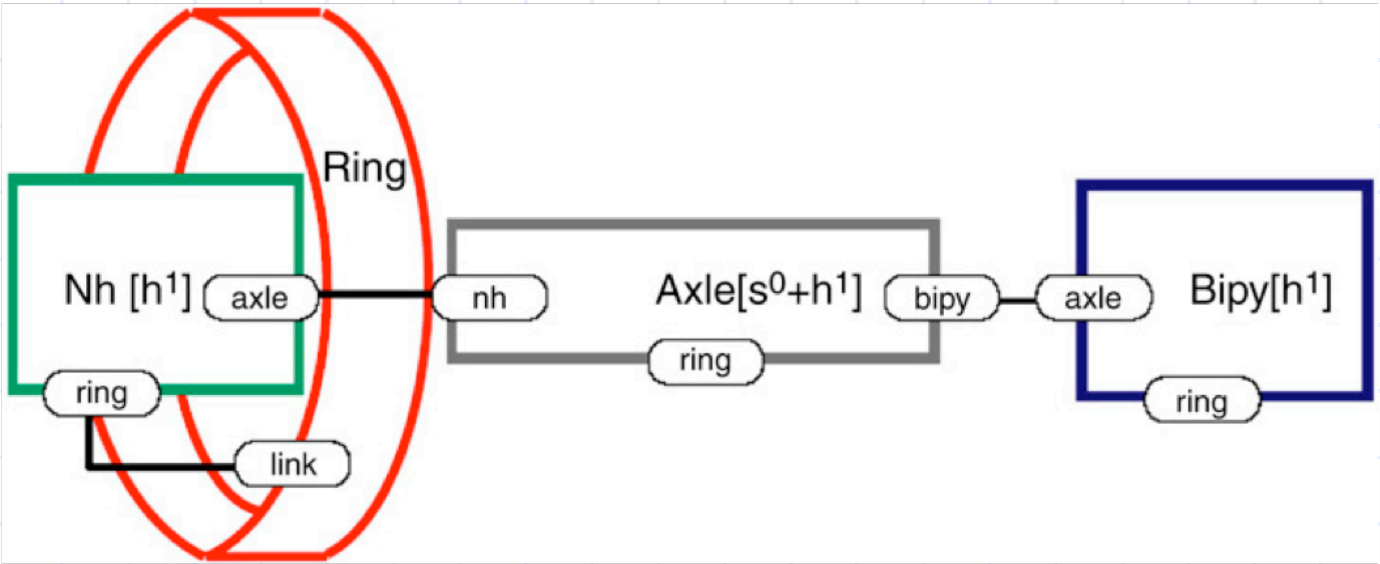
- Reactions are monotone:
  - either **creations** (bonds and/or proteins are created)
  - or **destructions** (bonds and/or proteins are removed)
  - or (extension of k) **bond flippings** (an edge of a bond may be moved from one molecule to another)



suggested by the bipedal movement on a DNA strand



RaH modeled in K: the initial state



$$Nh[h^1](axle^s + ring^x),\qquad Axle[s^0 + h^1](nh^s + bipy^r + ring),$$
$$Bipy[h^1](axle^r + ring),\qquad Ring\,(link^x)$$

RaH modeled in K: some reaction rules

- Deprotonation:

$$Nh[h^1], Base[h^0] \underset{base\_nh}{\overset{nh\_base}{\rightleftarrows}} Nh[h^0], Base[h^1]$$

- Disconnection:

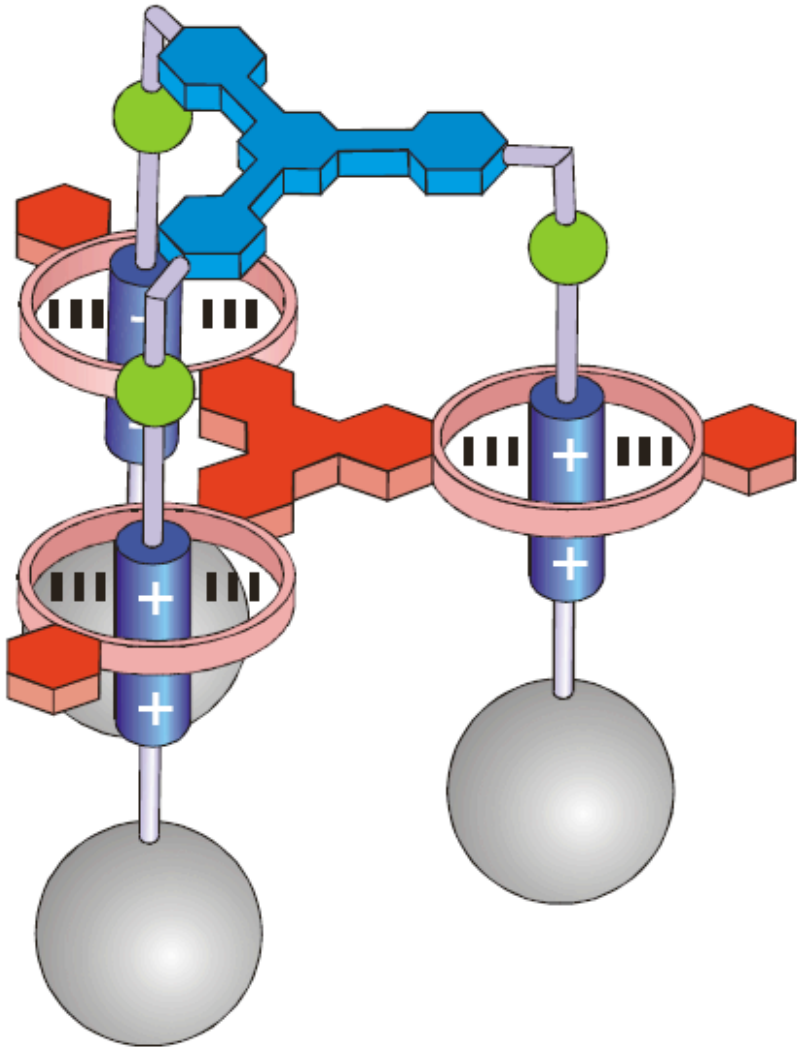
$$Nh[h^0](axle^s + ring^x), Axle[s^0](nh^s + ring) \overset{\infty}{\rightarrow} Nh[h^0](axle^s + ring), Axle[s^0](nh^s + ring^x)$$

- Connection:

$$Axle[s^0](bipy^r + ring^x), Bipy[h^0](axle^r + ring) \underset{unlink\_bipy}{\overset{link\_bipy}{\rightleftarrows}} Axle[s^1](bipy^r + ring), Bipy[h^0](axle^r + ring^x)$$

From the rotaxane to the nano-elevator

- **Nano-elevator**: three joined rotaxanes with a platform connected to the rings
- We plan to **re-use** the modeling of the rotaxane also for the nano-elevator (adding two new species: top & platform)
- **Monolithic approach**: exponential explosion of the species: $\mathbf{4^3 = 64}$

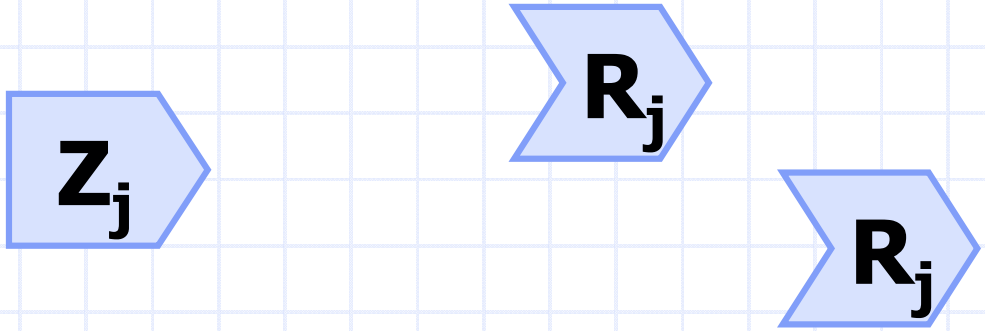


Plan of the talk

- Chemistry as a Computational Model
  - ... not a new issue
- Chemical Ground Form (CGF) [Cardelli08]
  - ... a new way to analyze chemistry
- Computational Strength of Chemistry [Concur08]
  - Non Turing universal but...
  - Turing complete formalisms can be approximated with any degree of precision
- From Chemistry to...
  - Biochemistry: BGF [AB08]
  - Nano devices: nano-K [CMSB07,TCS08]
- **Computational strength of BGF and K-calculi** [CMSB09]
- Concluding remarks

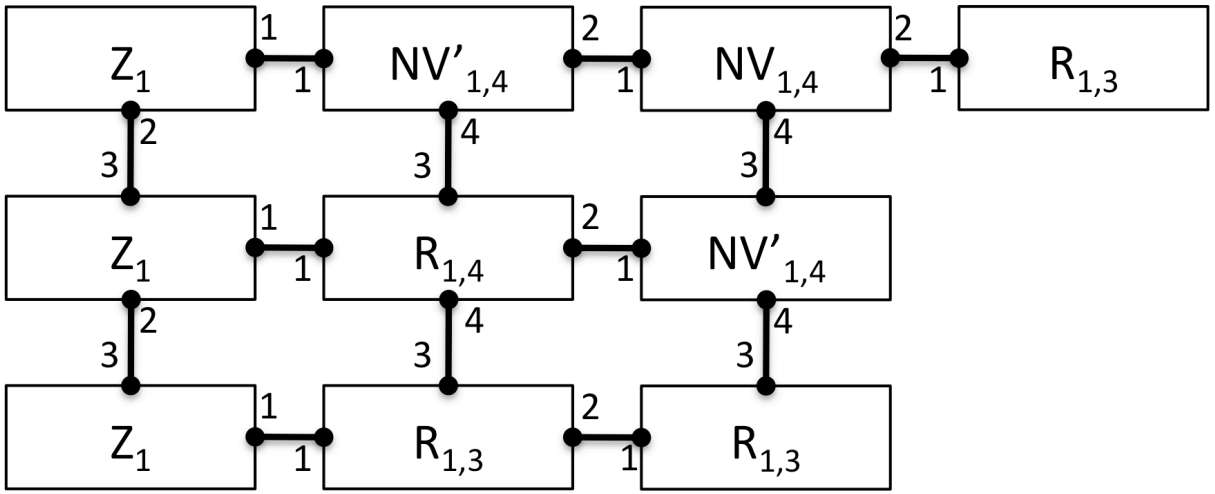
Turing completeness of BGF

- BGF=CGF+association+dissociation
  - This calculus is Turing complete
- **Proof**: RAMs can be faithfully modeled representing registers as polymers



Turing completeness of K-calculi

- K without **destructions** and **internal state** is already Turing complete
- **Proof**: RAMs can be faithfully modeled representing registers as grids



Plan of the talk

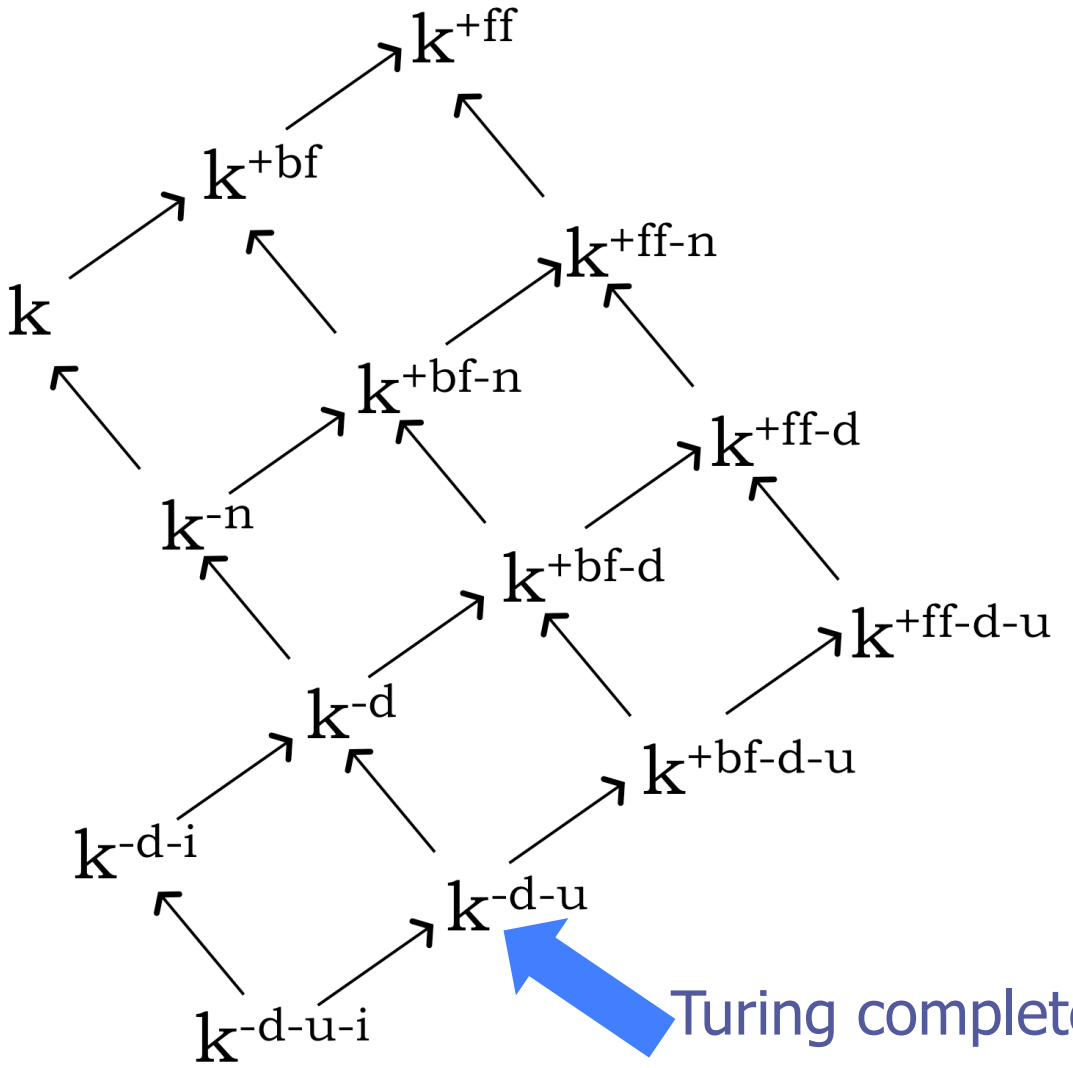
- Chemistry as a Computational Model
  - ... not a new issue
- Chemical Ground Form (CGF) [Cardelli08]
  - ... a new way to analyze chemistry
- Computational Strength of Chemistry [Concur08]
  - Non Turing universal but...
  - Turing complete formalisms can be approximated with any degree of precision
- From Chemistry to...
  - Biochemistry: BGF [AB08]
  - Nano devices: nano-K [CMSB07,TCS08]
- Computational strength of BGF and K-calculi [CMSB09]
- **Concluding remarks**

Conclusion

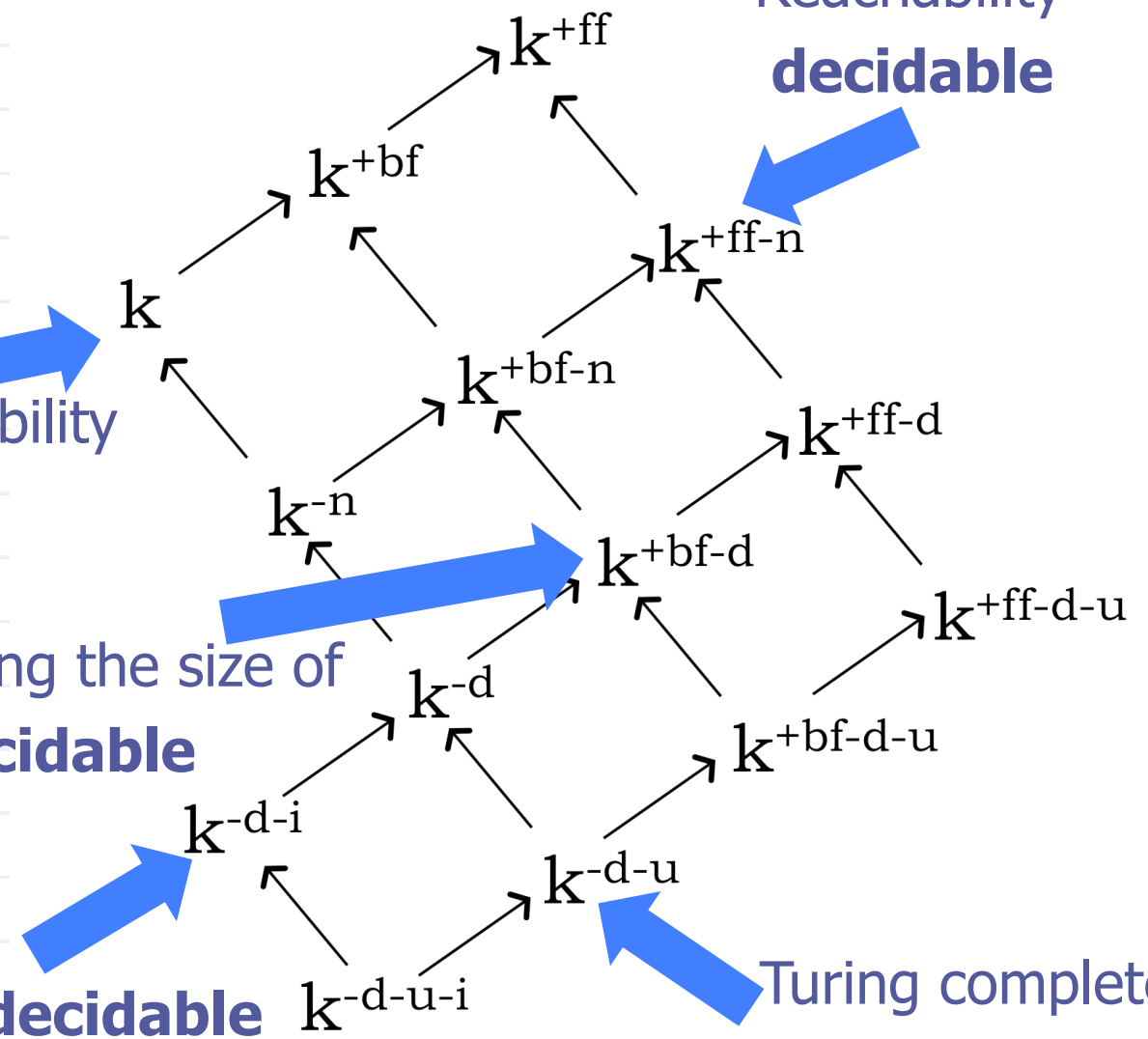
- CGF, the calculus for chemistry, is not Turing powerful…
  - …but Turing complete formalisms can be **approximated** with any given degree of precision
- Molecule complexation "Turifies" chemistry
  - BGF includes a **basic** mechanism for association/ dissociation, which is also biologically **realistic**, that makes CGF Turing complete

K-calculi: the lattice

- The ordering is the subcalculus relation
- Dialects are inspired by **biological** phenomena



K-calculi: the lattice



Reachability **decidable**

Reachability/Coverability **undecidable**

Coverability (limiting the size of complexes) **decidable**

Coverability **decidable**

Turing complete

Coverability **undecidable**

Future work

- Understand better the actual importance of molecule dissociation/bond destruction
  - K without **bond destruction** is Turing complete while BGF without **dissocation** is not!
  - Possible motivations: **multimolecular reactions** of K, test of the **entire interface** in K, ...
- **Quantitative** analysis
  - The approximated Turing completeness of CGF suggests that quantitative analysis is **undecidable**
    (is a molecule/complex reachable with prob. greater than P?)
  - As done for qualitative analysis, we plan to capture **fragments** of these calculi in which also quantitive analysis is possible

References

- Cardelli. **On process rate semantics**. *Theoretical Computer Science* 391(3): 190-215. 2008
- Zavattaro and Cardelli. **Termination problems in chemical kinetics**. In Proc. *Concur'08*. LNCS 5201:477-491. 2008
- Cardelli and Zavattaro. **On the computational power of biochemistry**. In Proc. *AB'08*. LNCS 5147:65-80. 2008
- Credi, Garavelli, Laneve, Pradalier, Silvi, and Zavattaro. **nanoκ:A calculus for the modeling and simulation of nano devices**. *Theoretical Computer Science 408(1): 17-30. 2008* (Preliminary version in CMSB'07).
- Delzanno, Di Giusto, Gabbrielli, Laneve, Zavattaro. **The κ-Lattice: Decidability Boundaries for Qualitative Analysis in Biological Languages**. To appear in *CMSB'09*.