

MODEL CHECKING OF BIOCHEMICAL NETWORKS USING PETRI NETS

Monika Heiner

**Brandenburg University of Technology
Cottbus**

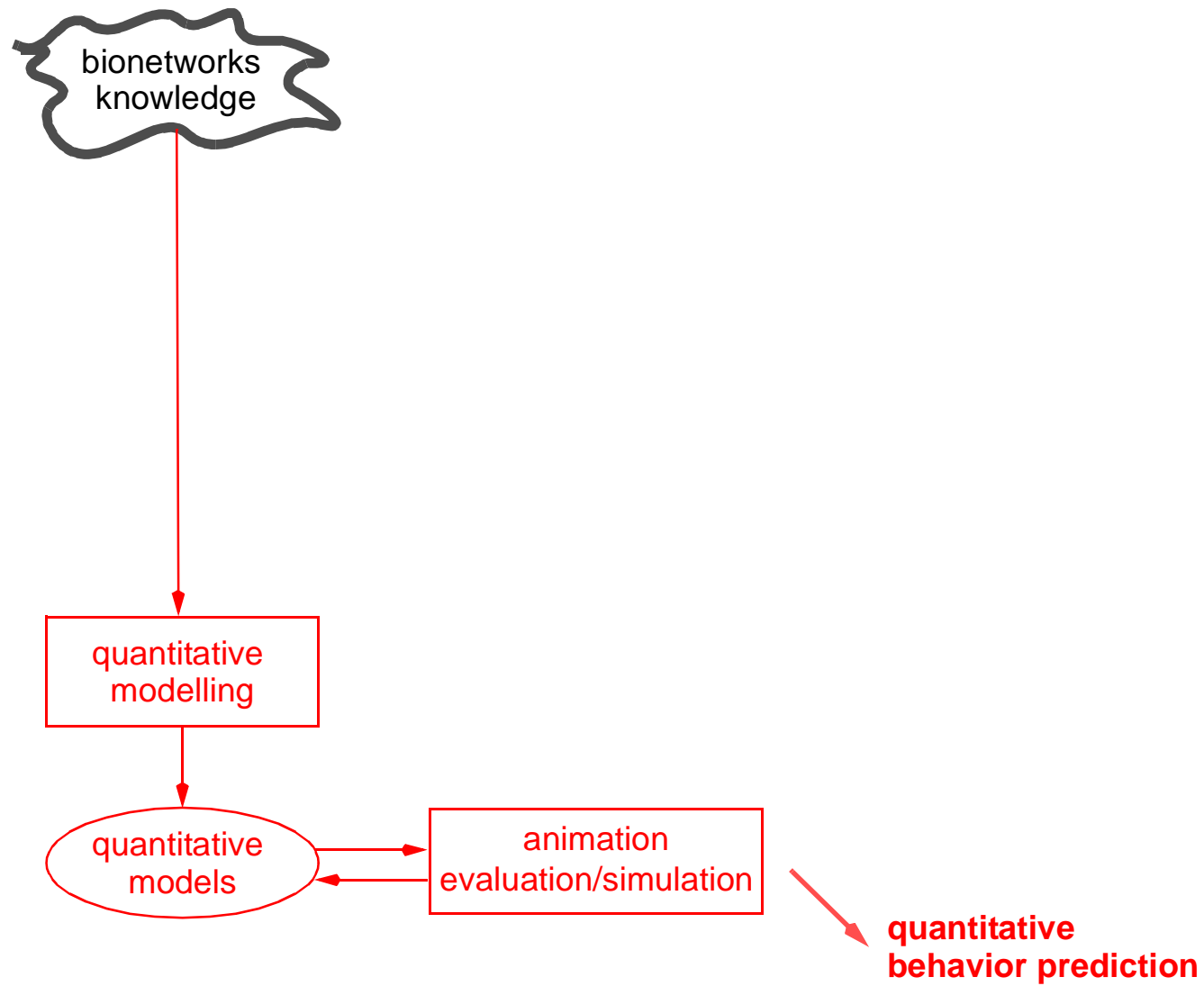
Dep. of CS

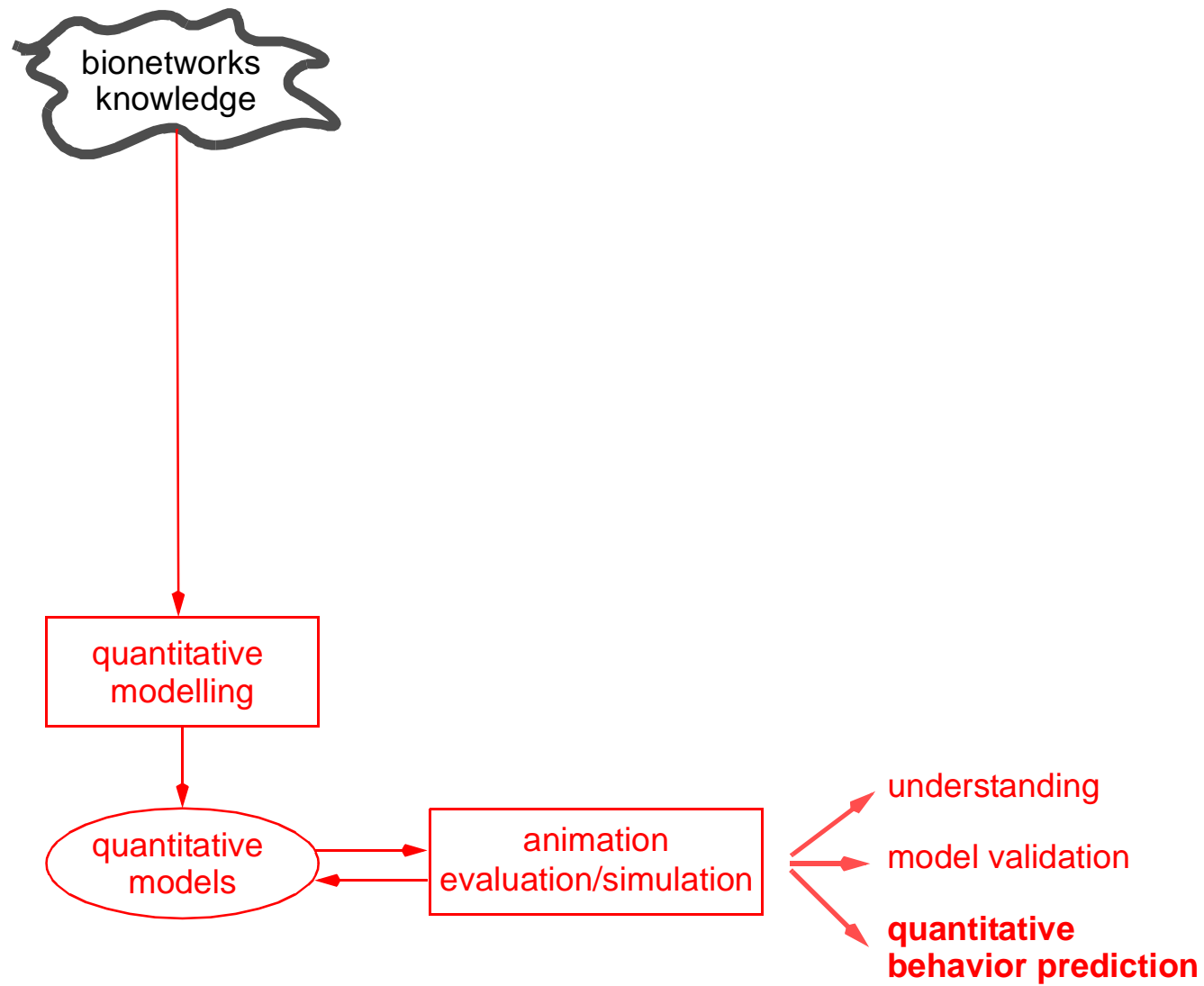
- ❑ **knowledge** **-> PROBLEM 1**
 - > *uncertain*
 - > *growing, changing*
 - > *distributed over independent data bases, papers, journals, . . .*

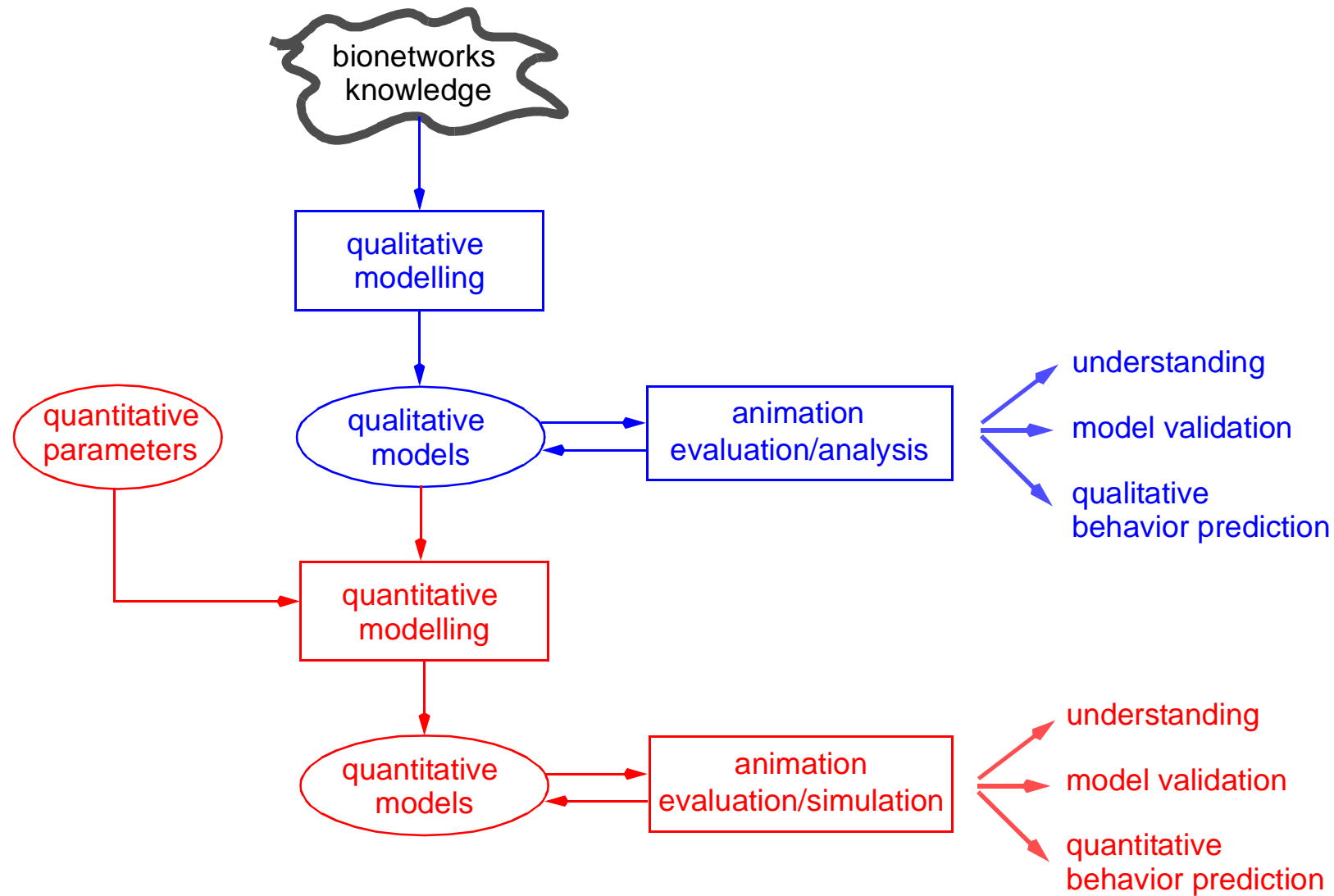
- ❑ **various, mostly ambiguous representations** **-> PROBLEM 2**
 - > *verbose descriptions*
 - > *diverse graphical representations*
 - > *contradictory and / or fuzzy statements*

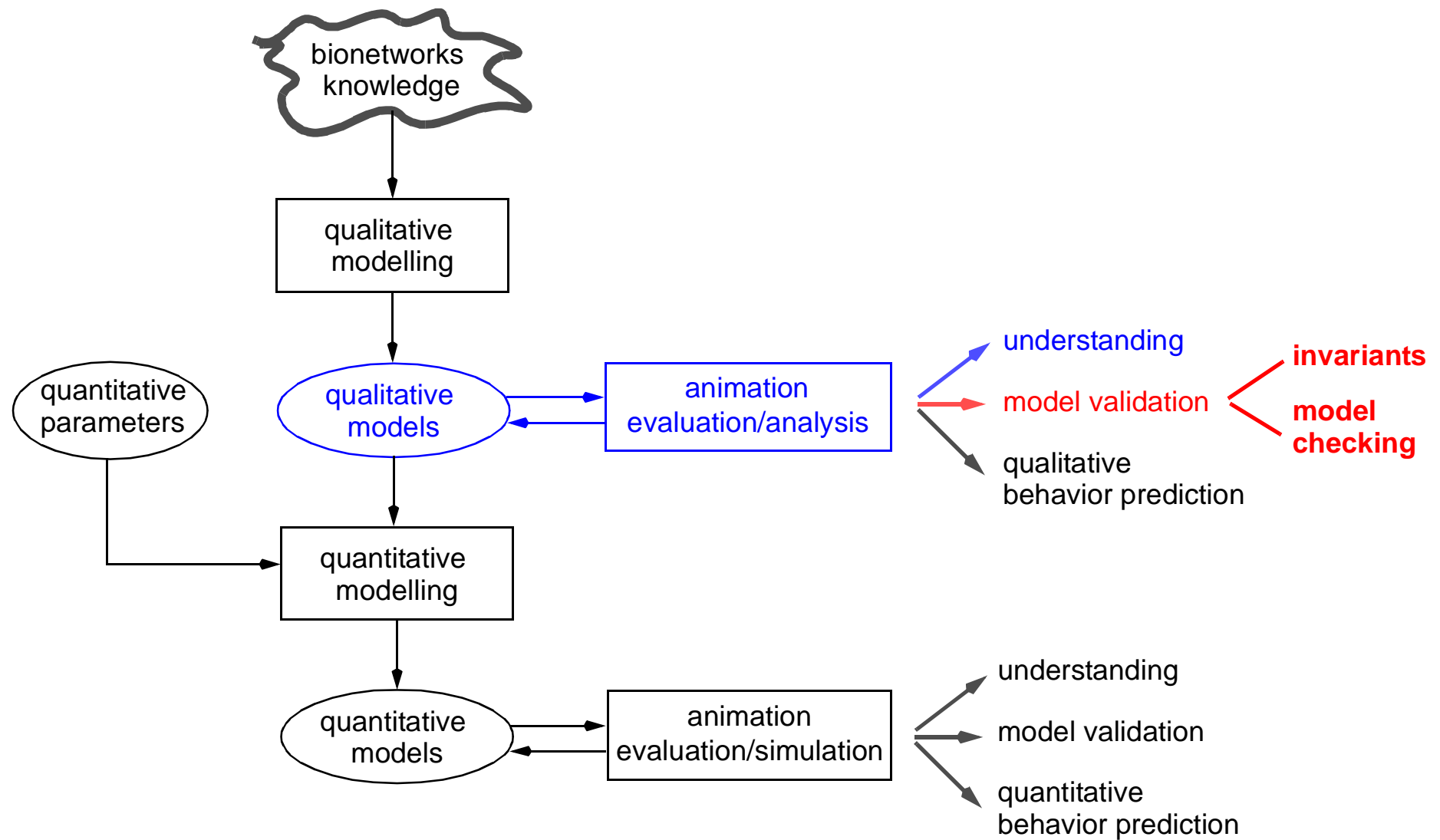
- ❑ **network structure** **-> PROBLEM 3**
 - > *tend to grow fast*
 - > *dense, apparently unstructured*
 - > *hard to read*

-->> models are full of ASSUMPTIONS <<--

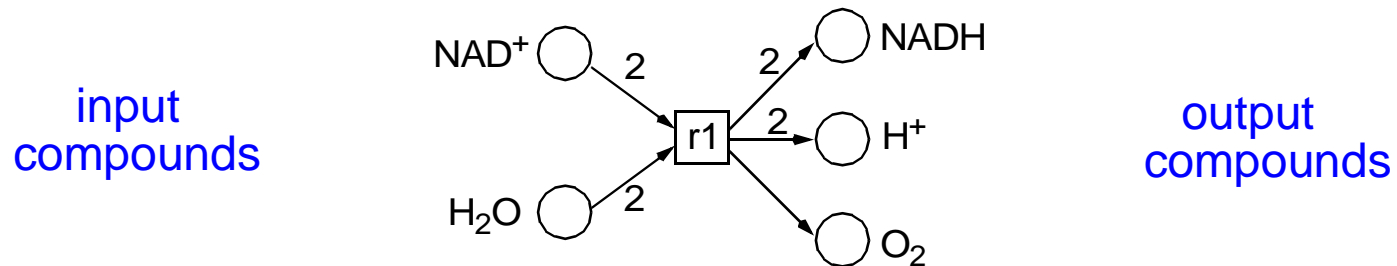
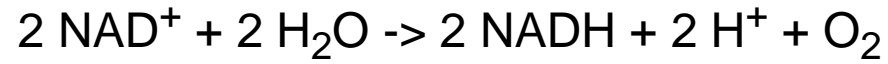




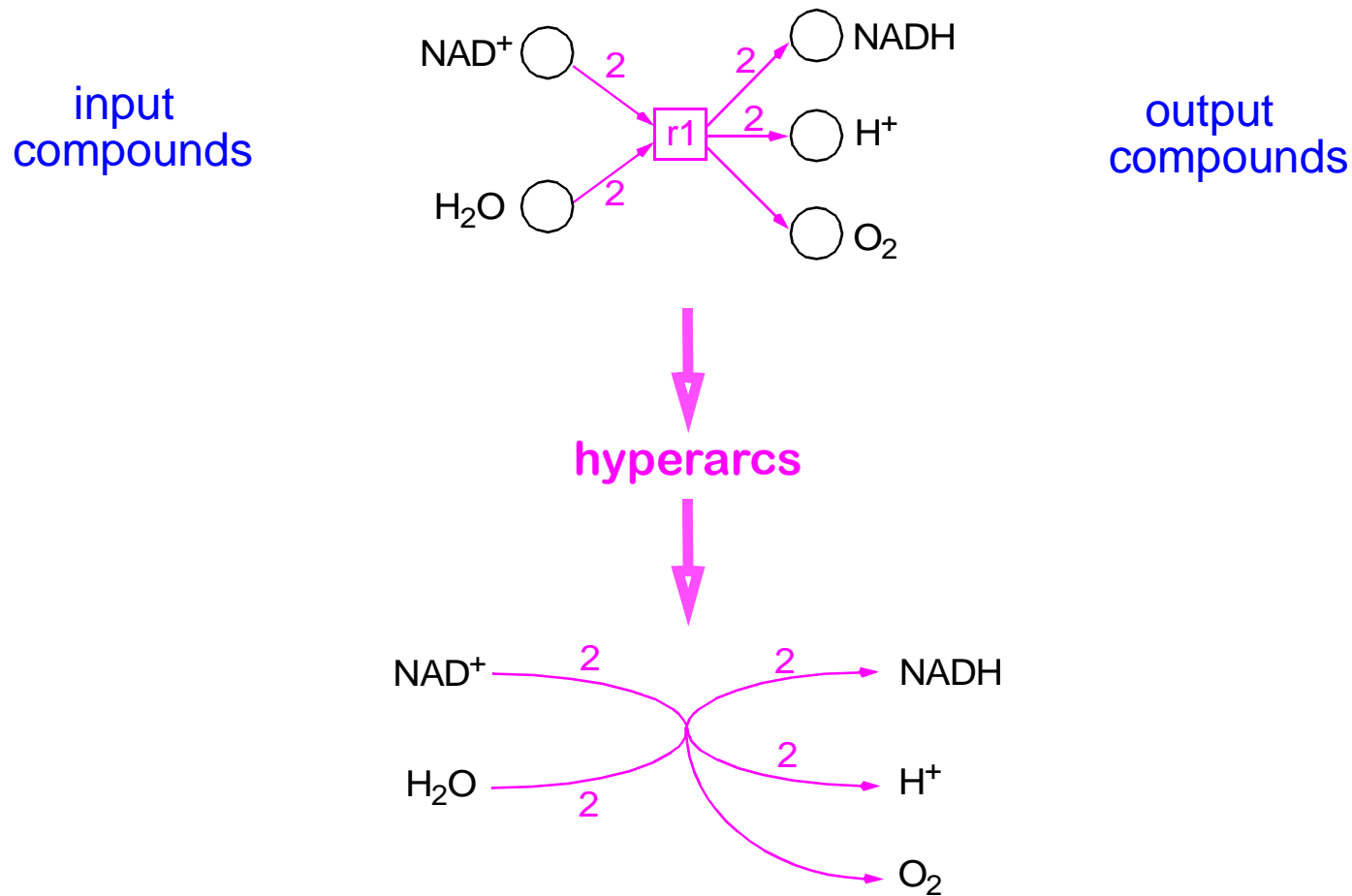
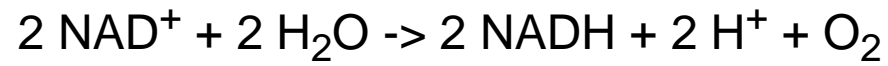




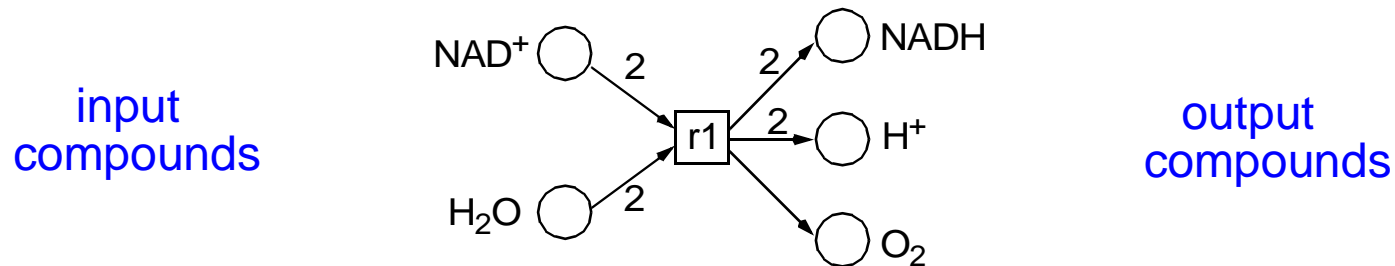
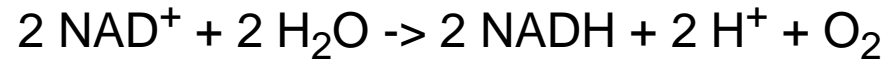
□ chemical reactions → atomic actions → Petri net transitions



□ chemical reactions → atomic actions → Petri net transitions

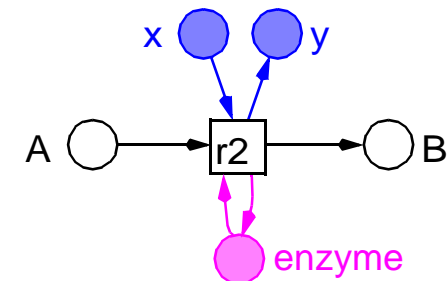


☐ chemical reactions → atomic actions → Petri net transitions

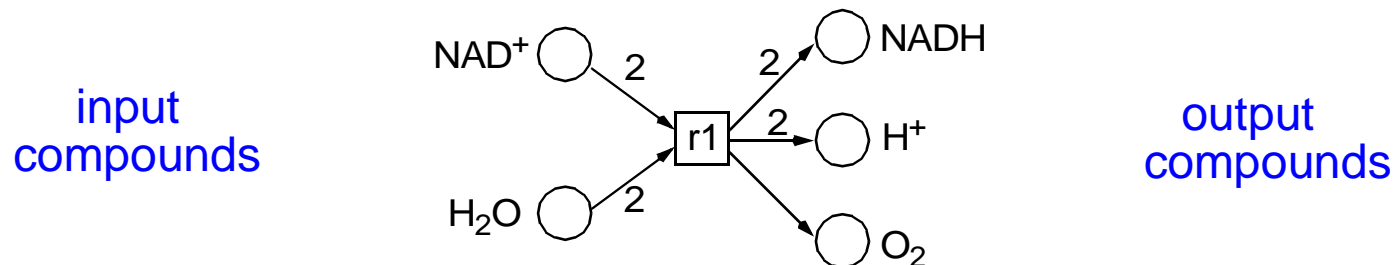
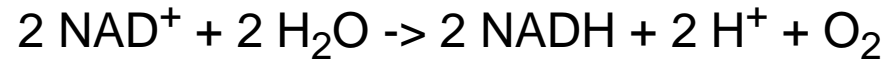


☐ chemical compounds → Petri net places

- | | |
|---|--|
| <ul style="list-style-type: none"> - primary compounds - auxiliary compounds, ubiquitous → fusion nodes - catalyzing compounds | <ul style="list-style-type: none"> - metabolites - e. g. electron carrier - enzymes |
|---|--|



☐ chemical reactions → atomic actions → Petri net transitions

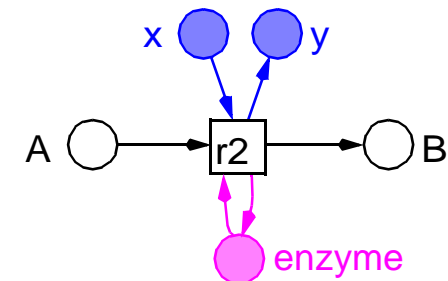


☐ chemical compounds

- primary compounds
- auxiliary compounds, ubiquitous → fusion nodes
- catalyzing compounds

→ Petri net places

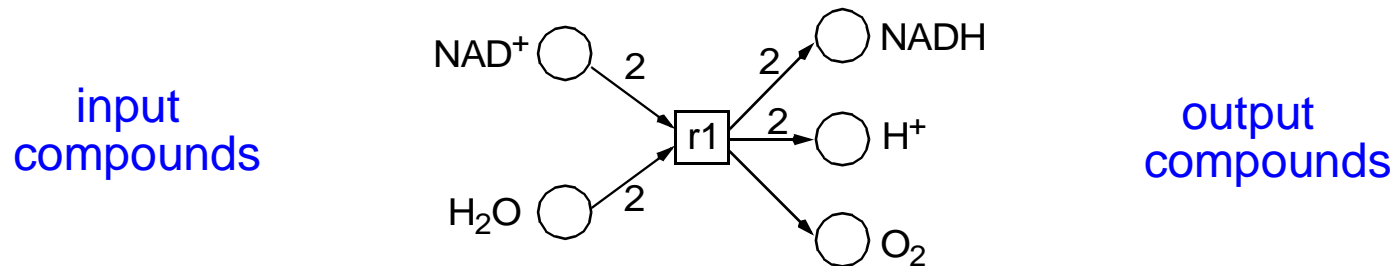
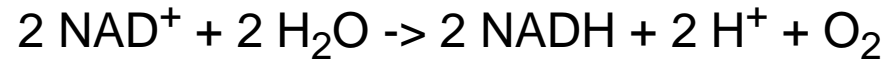
- metabolites
- e. g. electron carrier
- enzymes



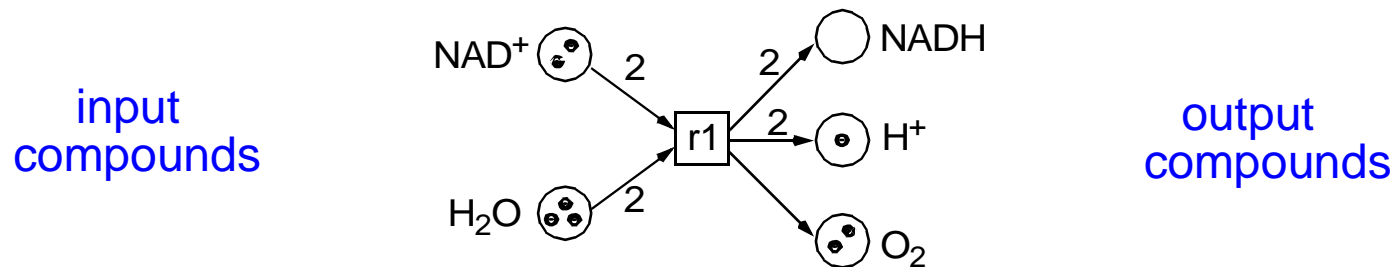
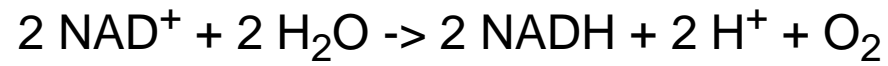
☐ stoichiometric relations → Petri net arc multiplicities

☐ compounds distribution → marking → tokens residing in places

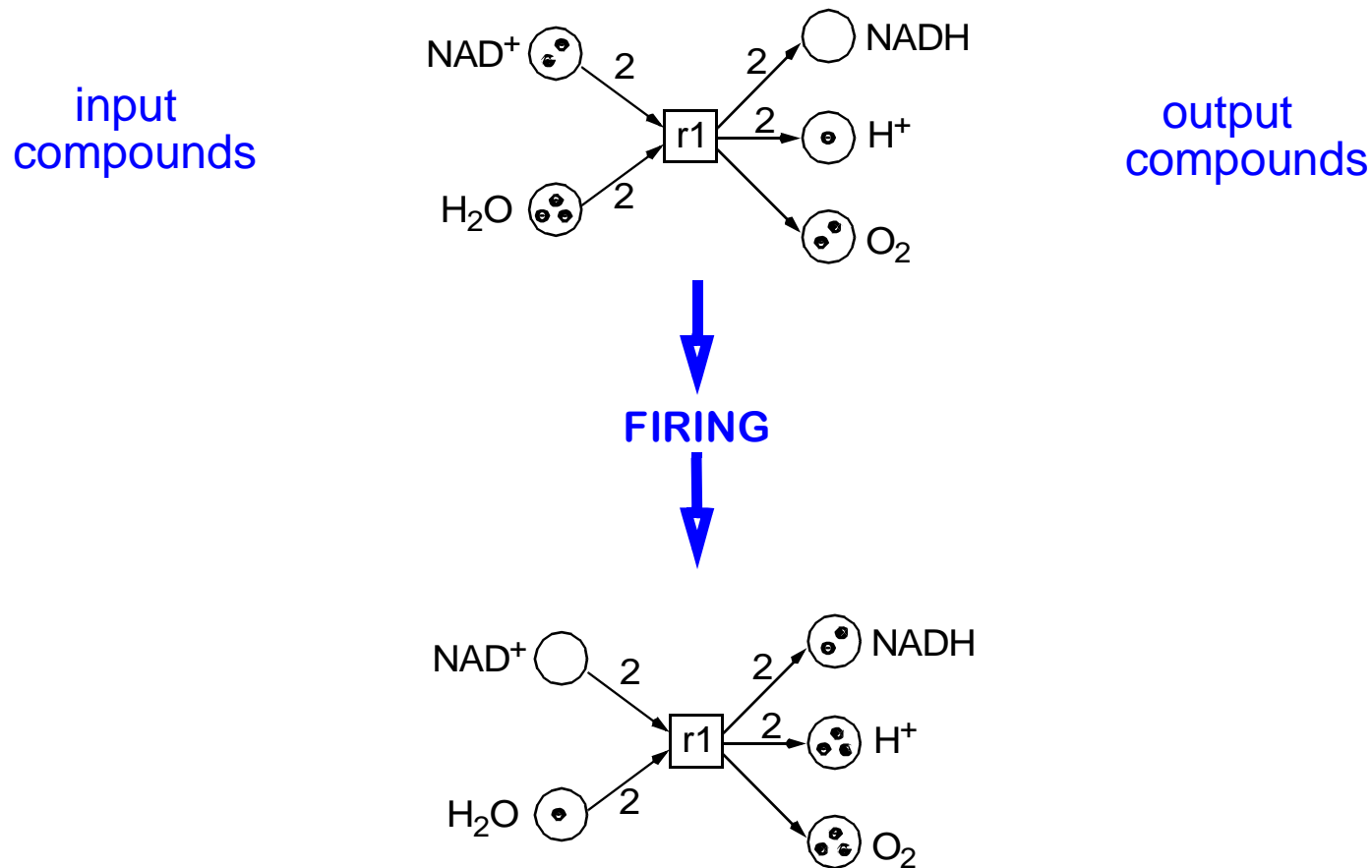
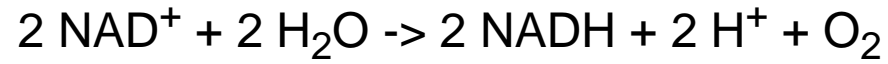
□ chemical reactions → atomic actions → Petri net transitions



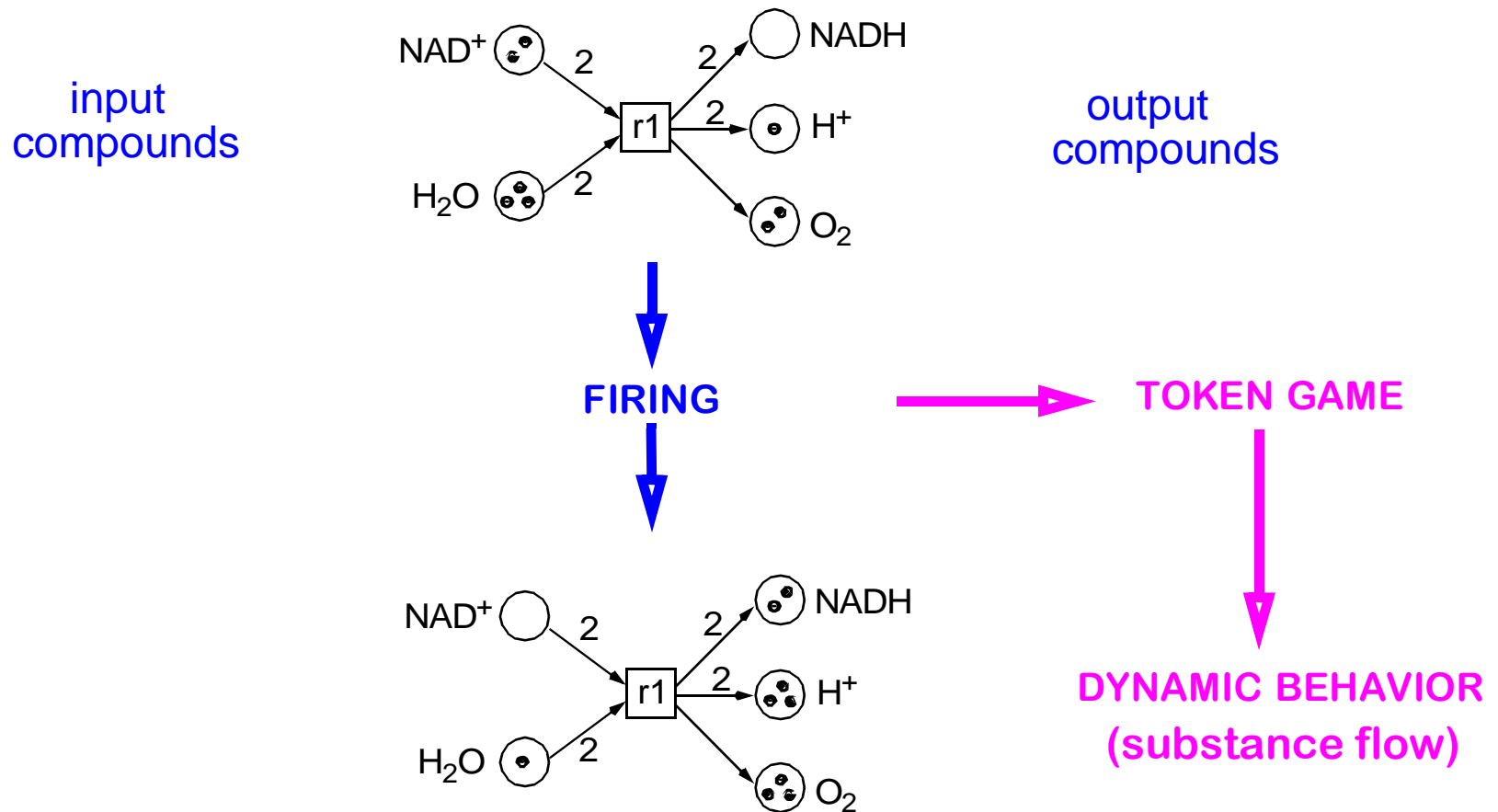
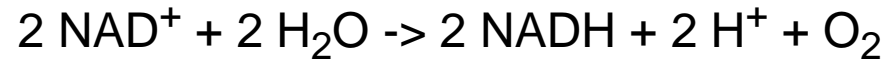
□ chemical reactions → atomic actions → Petri net transitions



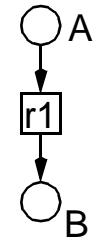
□ chemical reactions → atomic actions → Petri net transitions



□ chemical reactions → atomic actions → Petri net transitions



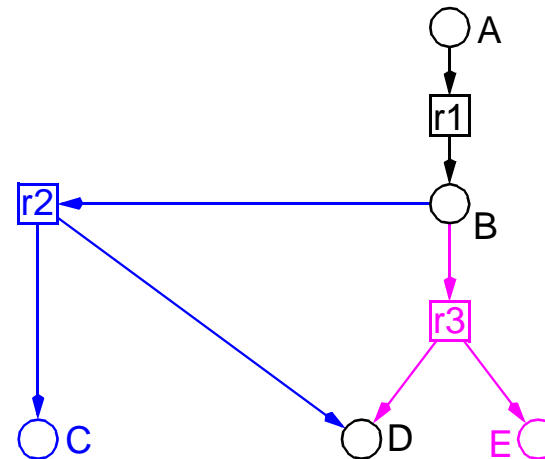
r1: A -> B



r1: A \rightarrow B

r2: B \rightarrow C + D

r3: B \rightarrow D + E



-> alternative reactions

r1: $A \rightarrow B$

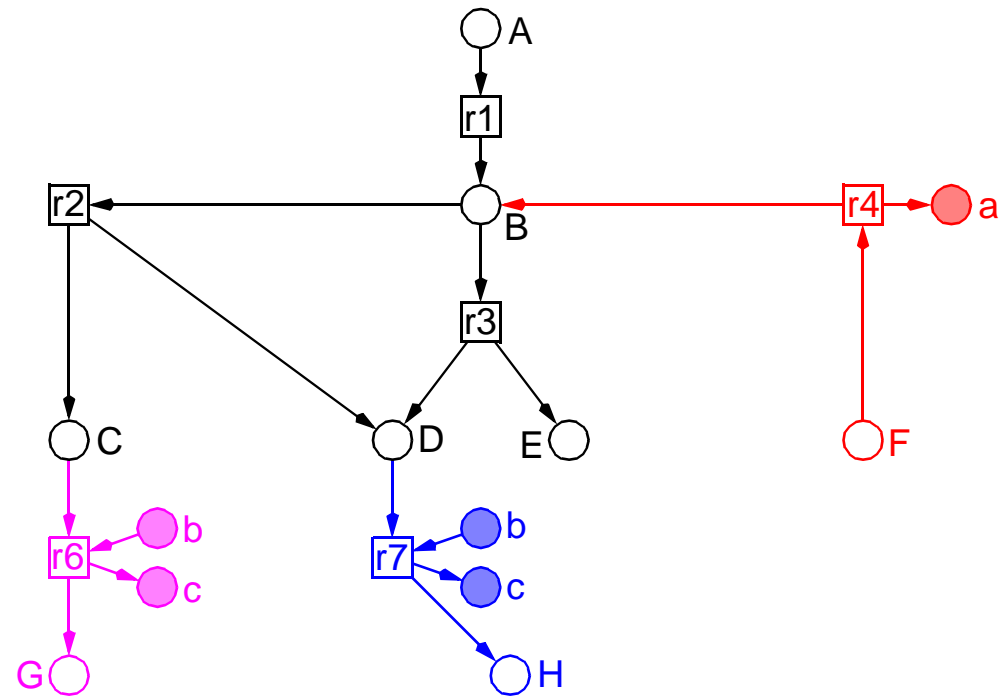
r2: $B \rightarrow C + D$

r3: $B \rightarrow D + E$

r4: $F \rightarrow B + a$

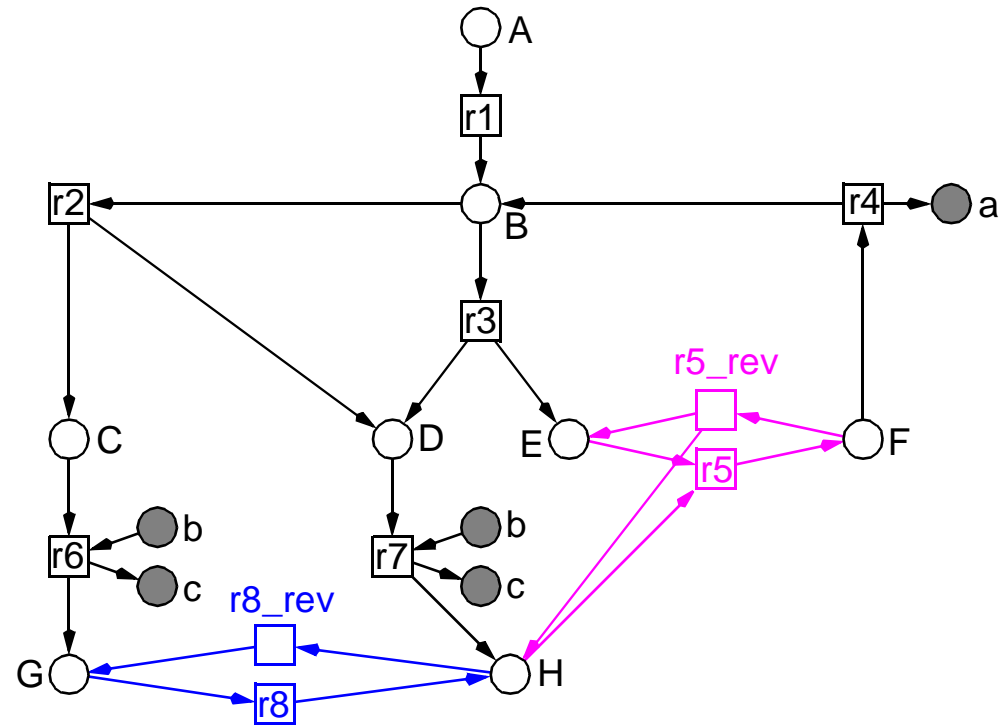
r6: $C + b \rightarrow G + c$

r7: $D + b \rightarrow H + c$



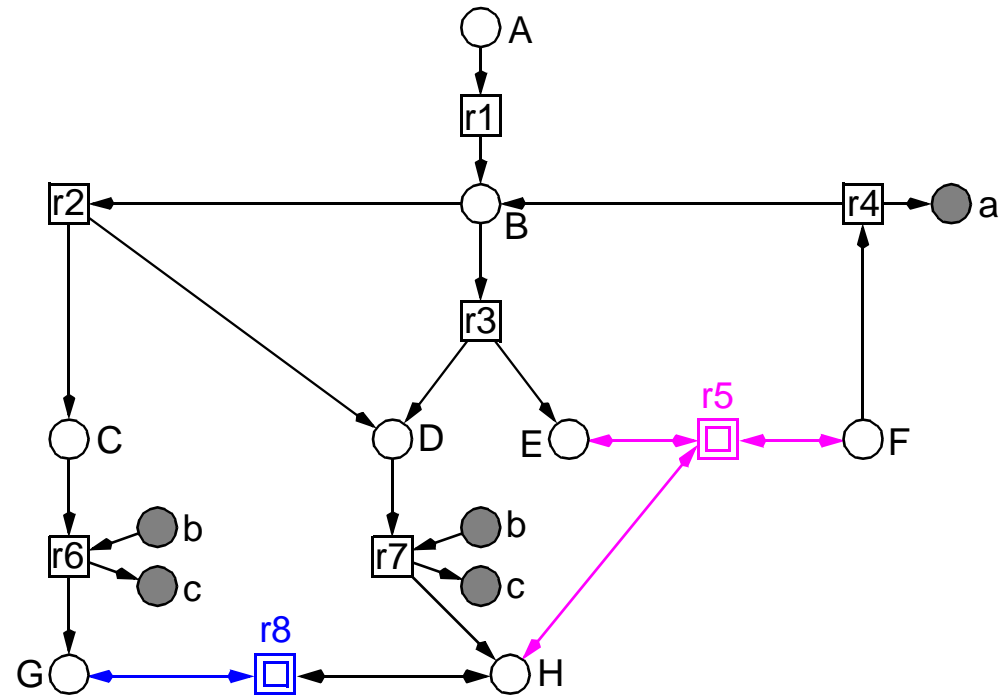
-> concurrent reactions

- r1: $A \rightarrow B$
- r2: $B \rightarrow C + D$
- r3: $B \rightarrow D + E$
- r4: $F \rightarrow B + a$
- r5: $E + H \leftrightarrow F$
- r6: $C + b \rightarrow G + c$
- r7: $D + b \rightarrow H + c$
- r8: $H \leftrightarrow G$



-> reversible reactions

- r1: $A \rightarrow B$
- r2: $B \rightarrow C + D$
- r3: $B \rightarrow D + E$
- r4: $F \rightarrow B + a$
- r5: $E + H \leftrightarrow F$
- r6: $C + b \rightarrow G + c$
- r7: $D + b \rightarrow H + c$
- r8: $H \leftrightarrow G$



-> reversible reactions
- hierarchical nodes

r1: $A \rightarrow B$

r2: $B \rightarrow C + D$

r3: $B \rightarrow D + E$

r4: $F \rightarrow B + a$

r5: $E + H \leftrightarrow F$

r6: $C + b \rightarrow G + c$

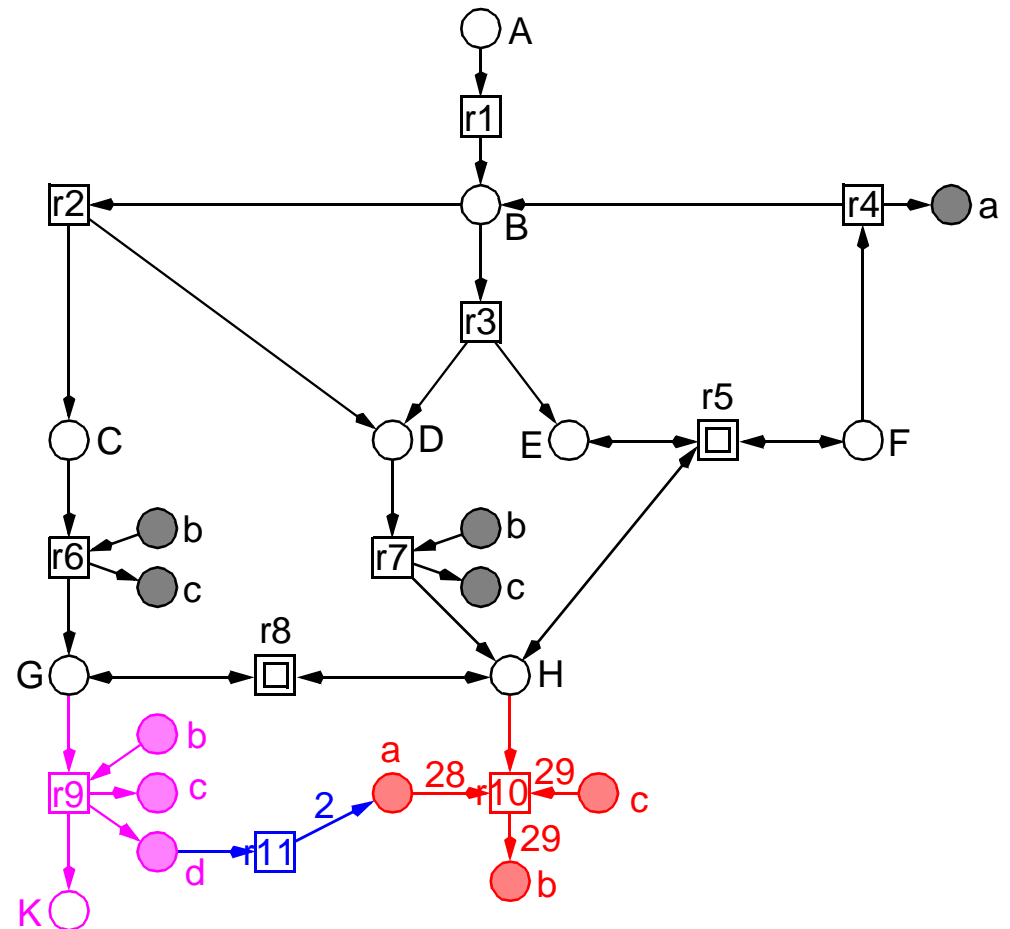
r7: $D + b \rightarrow H + c$

r8: $H \leftrightarrow G$

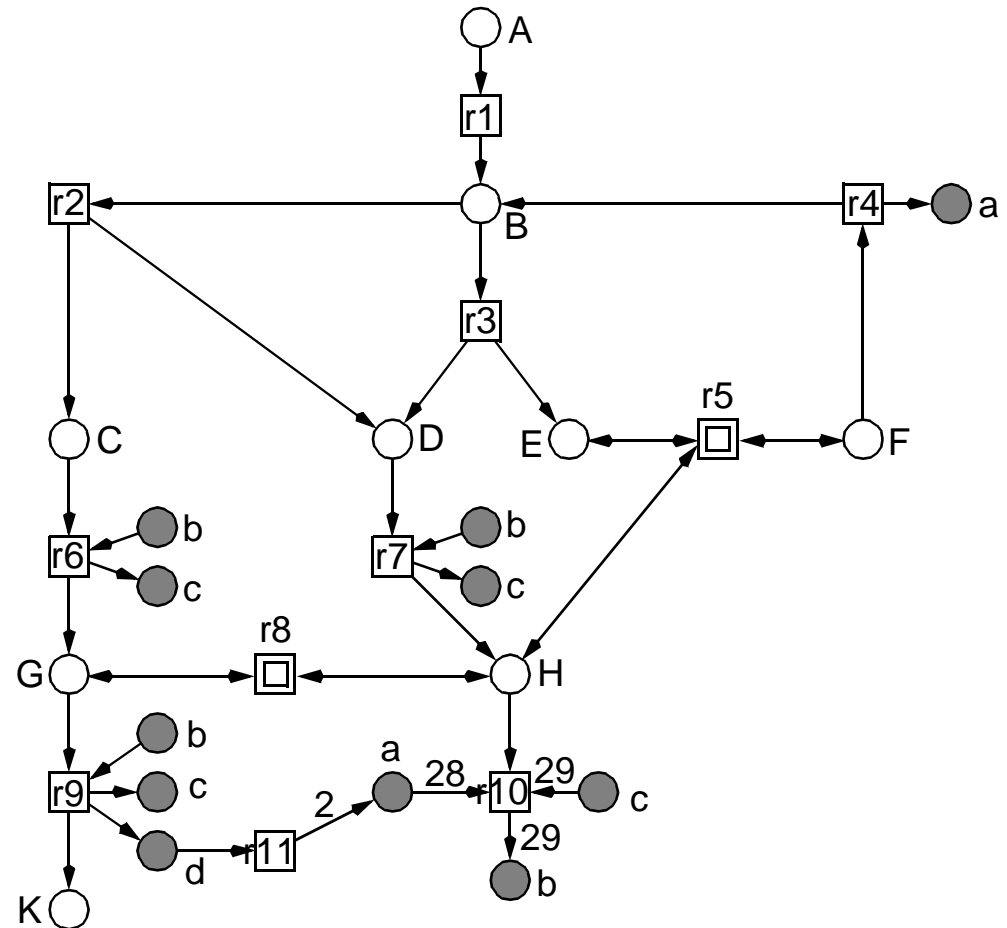
r9: $G + b \rightarrow K + c + d$

r10: $H + 28a + 29c \rightarrow 29b$

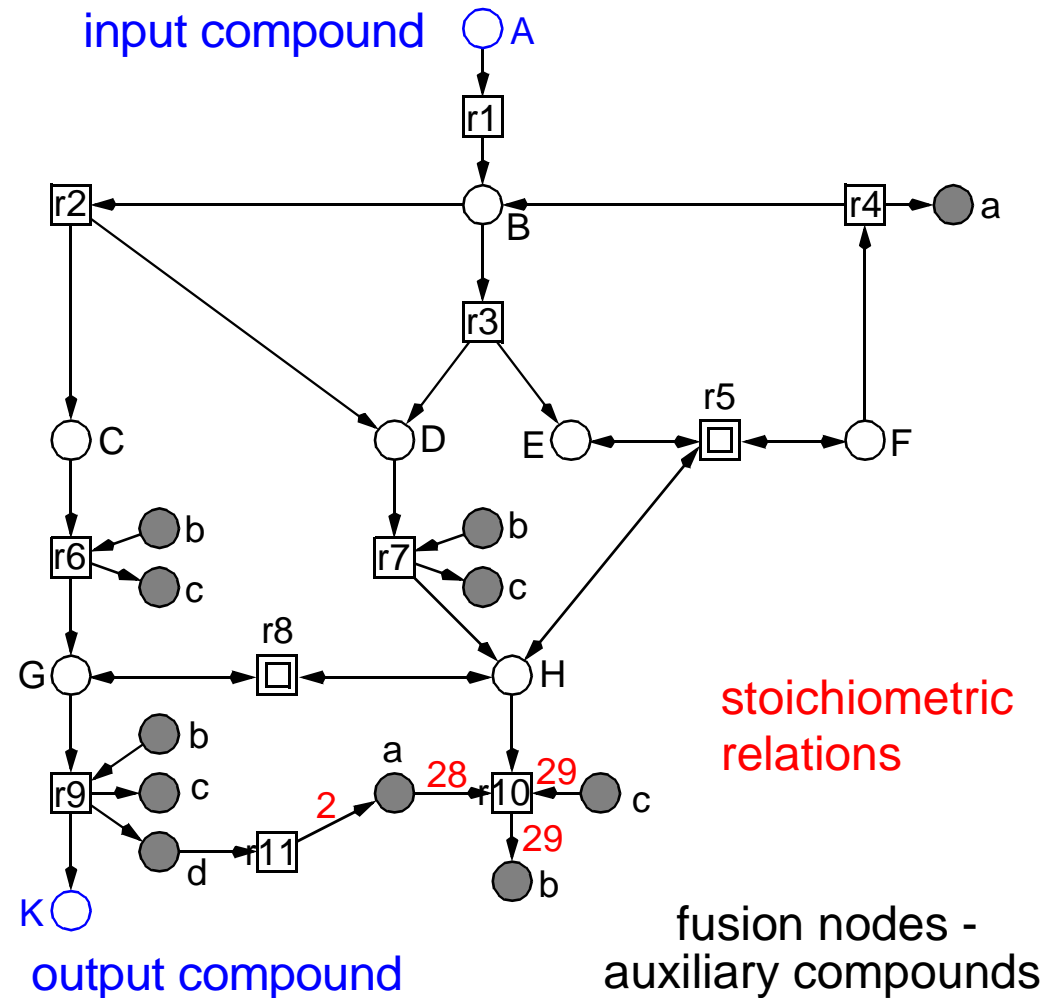
r11: $d \rightarrow 2a$



- r1: $A \rightarrow B$
- r2: $B \rightarrow C + D$
- r3: $B \rightarrow D + E$
- r4: $F \rightarrow B + a$
- r5: $E + H \leftrightarrow F$
- r6: $C + b \rightarrow G + c$
- r7: $D + b \rightarrow H + c$
- r8: $H \leftrightarrow G$
- r9: $G + b \rightarrow K + c + d$
- r10: $H + 28a + 29c \rightarrow 29b$
- r11: $d \rightarrow 2a$



- r1: $A \rightarrow B$
- r2: $B \rightarrow C + D$
- r3: $B \rightarrow D + E$
- r4: $F \rightarrow B + a$
- r5: $E + H \leftrightarrow F$
- r6: $C + b \rightarrow G + c$
- r7: $D + b \rightarrow H + c$
- r8: $H \leftrightarrow G$
- r9: $G + b \rightarrow K + c + d$
- r10: $H + 28a + 29c \rightarrow 29b$
- r11: $d \rightarrow 2a$



- ❑ networks of chemical reactions

- ❑ biologically interpreted Petri nets -
partial order sequences of chemical reactions
 - > *transforming input into output compounds*
 - > *respecting the given stoichiometric relations*

- ❑ networks of chemical reactions
- ❑ biologically interpreted Petri nets - **partial order sequences** of chemical reactions
 - > *transforming input into output compounds*
 - > *respecting the given stoichiometric relations*

❑ typical (structural) properties

INA

ORD	HOM	NBM	PUR	CSV	SCF	CON	SC	Ft0	tF0	Fp0	pF0	MG	SM	FC	EFC	ES
N	N	N	Y	N	N	Y	N	N	N	Y	Y	N	N	N	N	N
DTP	CPI	CTI	B	SB	REV	DSt	BSt	DTr	DCF	L	LV	L&S				
N	N	N	Y	Y	?	?	?	?	?	N	?	N				

- ❑ networks of chemical reactions
- ❑ biologically interpreted Petri nets - **partial order sequences** of chemical reactions
 - > transforming input into output compounds
 - > respecting the given stoichiometric relations

- ❑ typical (structural) properties

INA																			
ORD	HOM	NBM	PUR	CSV	SCF	CON	SC	Ft0	tF0	Fp0	pF0	MG	SM	FC	EFC	ES			
N	N	N	Y	N	N	Y	N	N	N	Y	Y	N	N	N	N	N			
DTP	CPI	CTI	B	SB	REV	DSt	BSt	DTr	DCF	L	LV	L&S							
N	N	N	Y	Y	?	?	?	?	?	N	?	N							

- ❑ **networks need environment behavior !!**

□ STYLE 1

- > *weak* assumptions
- > *infinite flow into/out the network*

□ STYLE 2

- > *firm* assumptions
- > *infinite many primary compounds*
- > *finite, but sufficient reservoir of auxiliary compounds*

□ STYLE 3

- > *strong* assumptions
- > *finite, but sufficient reservoir of auxiliary compounds*
- > *quantitative relations of input/output compounds*
 - *finite reservoir of primary compounds*

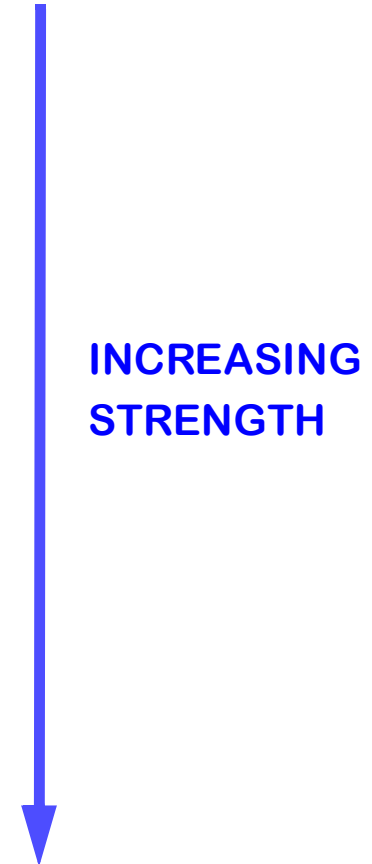


**INCREASING
STRENGTH**

- ❑ **STYLE 1** **-> VALIDATION CRITERION 1**
 - > *weak* assumptions
 - > *infinite flow into/out the network*

- ❑ **STYLE 2** **-> VALIDATION CRITERION 2**
 - > *firm* assumptions
 - > *infinite many primary compounds*
 - > *finite, but sufficient reservoir of auxiliary compounds*

- ❑ **STYLE 3** **-> VALIDATION CRITERION 3**
 - > *strong* assumptions
 - > *finite, but sufficient reservoir of auxiliary compounds*
 - > *quantitative relations of input/output compounds*
 - *finite reservoir of primary compounds*



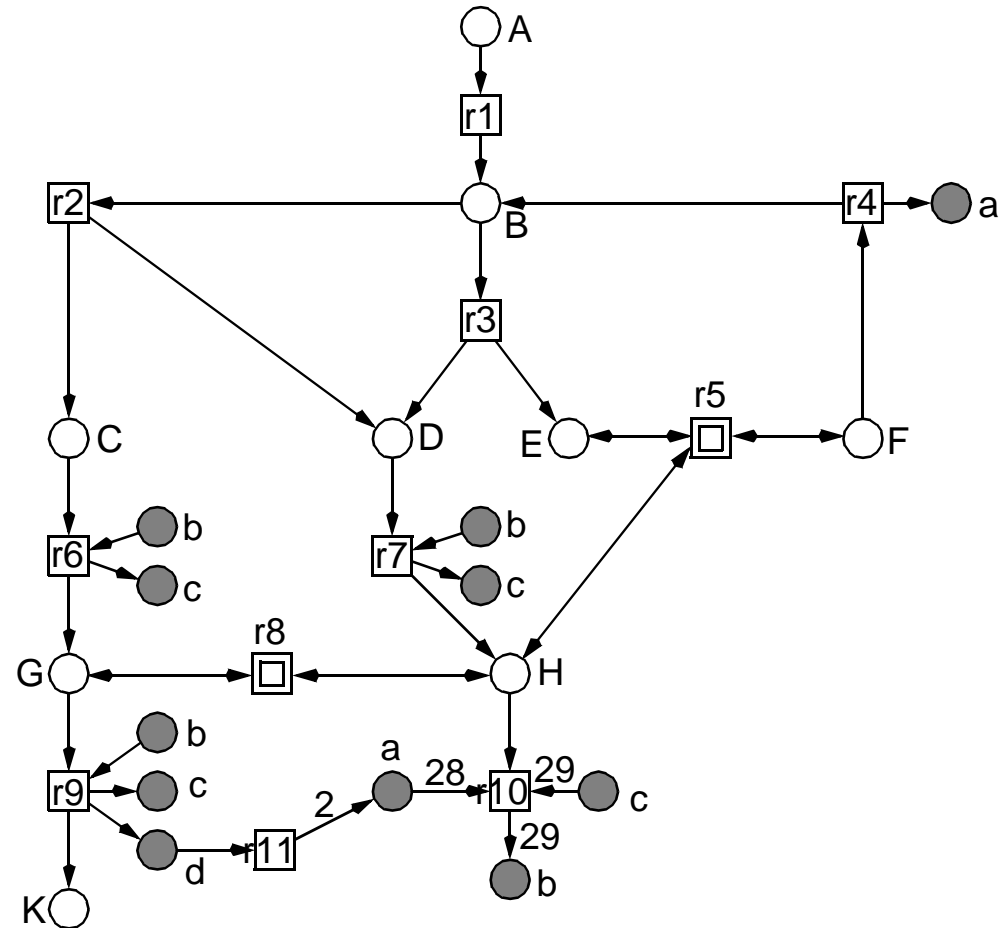
- ❑ **validation criteria ?**

- > *objective*
- > *reproducible*
- > *computable*

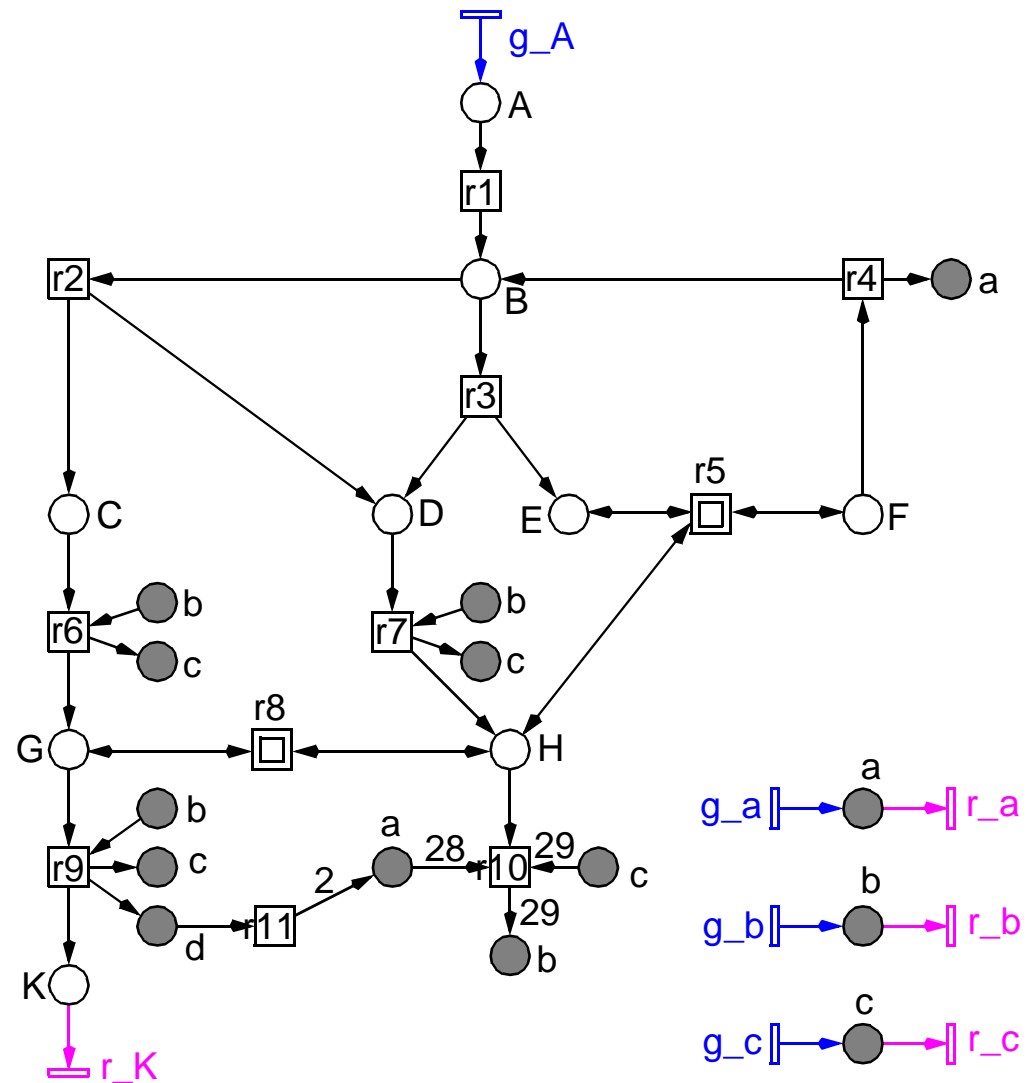
- ❑ **steady state behaviour**

- > *all flows preserving the given compounds distribution*

- ❑ **weak assumptions**
- ❑ **to animate the model**
 - > *deeper insights*
 - > *infinite substance flow*
- ❑ **to validate the model**



- ❑ **weak assumptions**
- ❑ **input substances**
-> *generating pre-transitions*
- ❑ **output substances**
-> *consuming post-transitions*
- ❑ **auxiliary substances**
-> *both*
- ❑ **no boundary places, but boundary transitions**
- ❑ **transitions without pre-places**
-> *live*
-> *all post-places are unbounded*

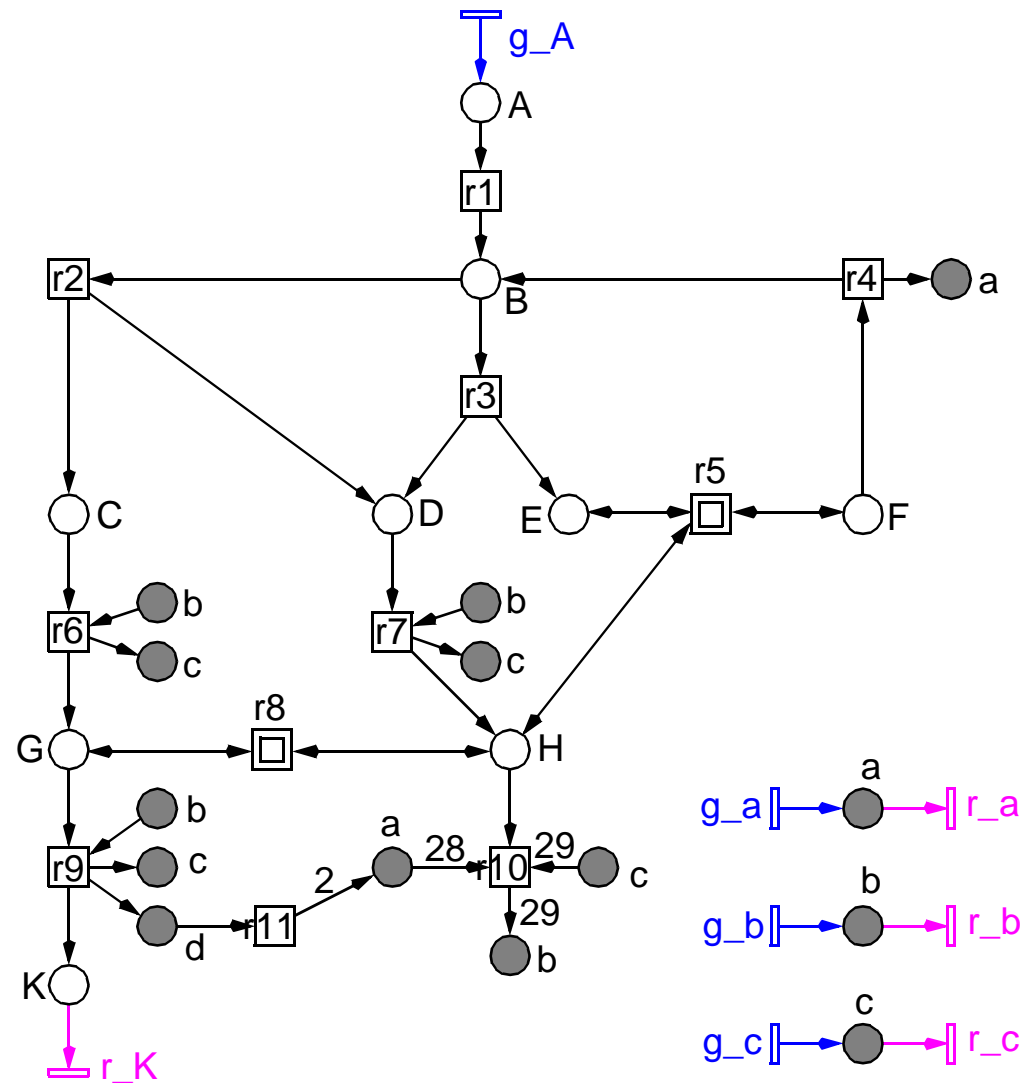


- **weak assumptions**

- **initial marking**
-> *empty marking*

- **steady state behaviour**
-> *empty marking reproduction*

- **(cause effect) pathway**
-> *minimal flow preserving given compounds distribution*
-> *elementary modes*
= *minimal T-invariants*



- ❑ **Lautenbach, 1973**
- ❑ **T-invariants** -> *multisets of transitions*
 - > *integer solutions of $Cx = 0, x \neq 0, x \geq 0$*
- ❑ **minimal T-invariants**
 - > *there is no T-invariant with a smaller support* -> *sets of transitions*
 - > *gcd of all entries is 1*
- ❑ **any T-invariant is a non-negative linear combination of minimal ones**
 - > *multiplication with a positive integer*
 - > *addition*
 - > *Division by gcd*
- ❑ **Covered by T-Invariants (CTI)**
 - > *each transition belongs to a T-invariant*

- ❑ Lautenbach, 1973
- ❑ T-invariants
 - > *integer* solutions of $Cx = 0, x \neq 0, x \geq 0$ -> *multisets of transitions*
 - > *exponential complexity*
- ❑ minimal T-invariants
 - > *there is no T-invariant with a smaller support* -> *sets of transitions*
 - > *gcd of all entries is 1*
- ❑ any T-invariant is a non-negative linear combination of minimal ones
 - > *multiplication with a positive integer*
 - > *addition*
 - > *Division by gcd*
- ❑ Covered by T-Invariants (CTI)
 - > *each transition belongs to a T-invariant*

trivial min. T-invariants (5)

- boundary transitions of auxiliary compounds

-> (g_a, r_a) , (g_b, r_b) ,
 (g_c, r_c)

- reversible reactions

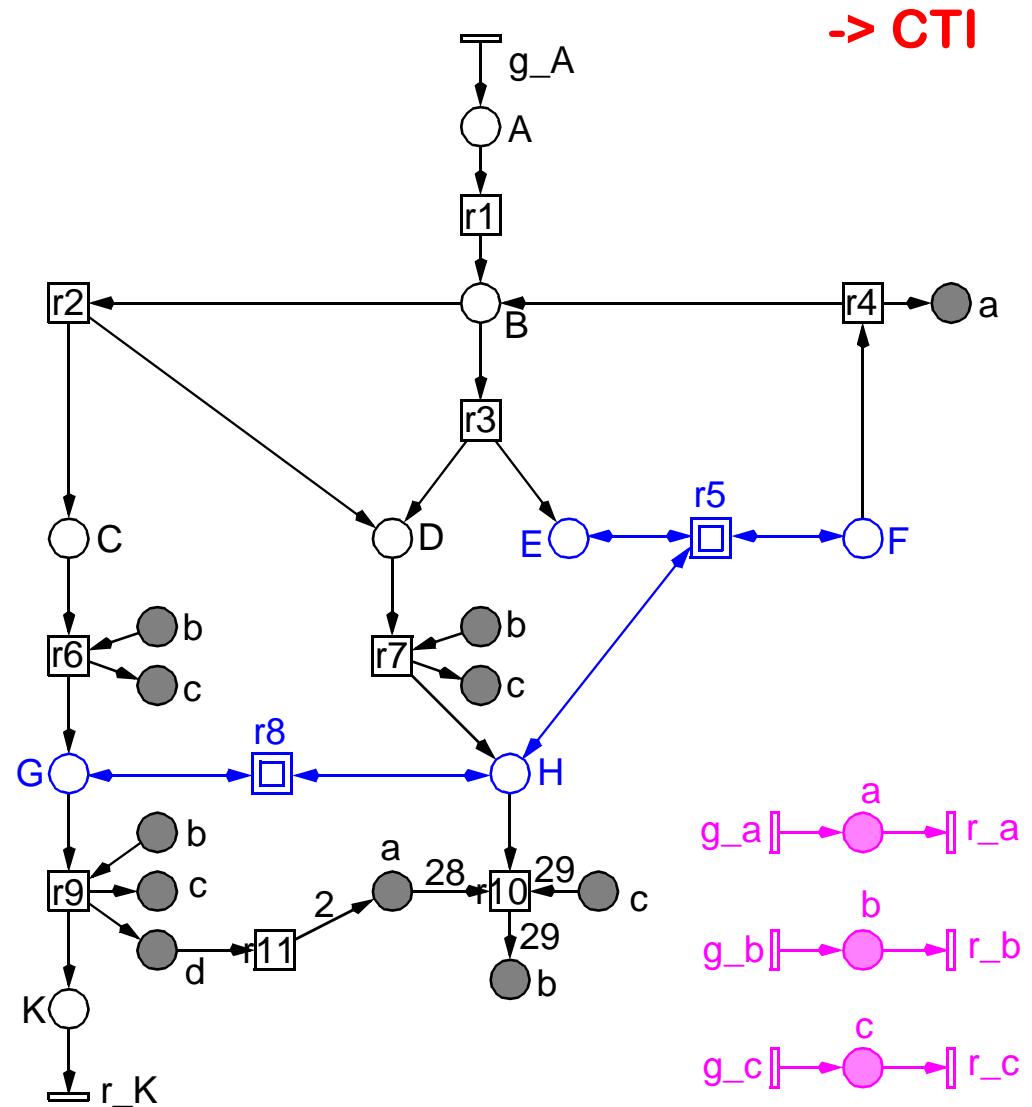
-> $(r5, r5_{rev})$, $(r8, r8_{rev})$

non-trivial min. T-invariants (7)

- covering boundary transitions of input / output compounds

-> *i/o-T-invariants*

- inner cycles

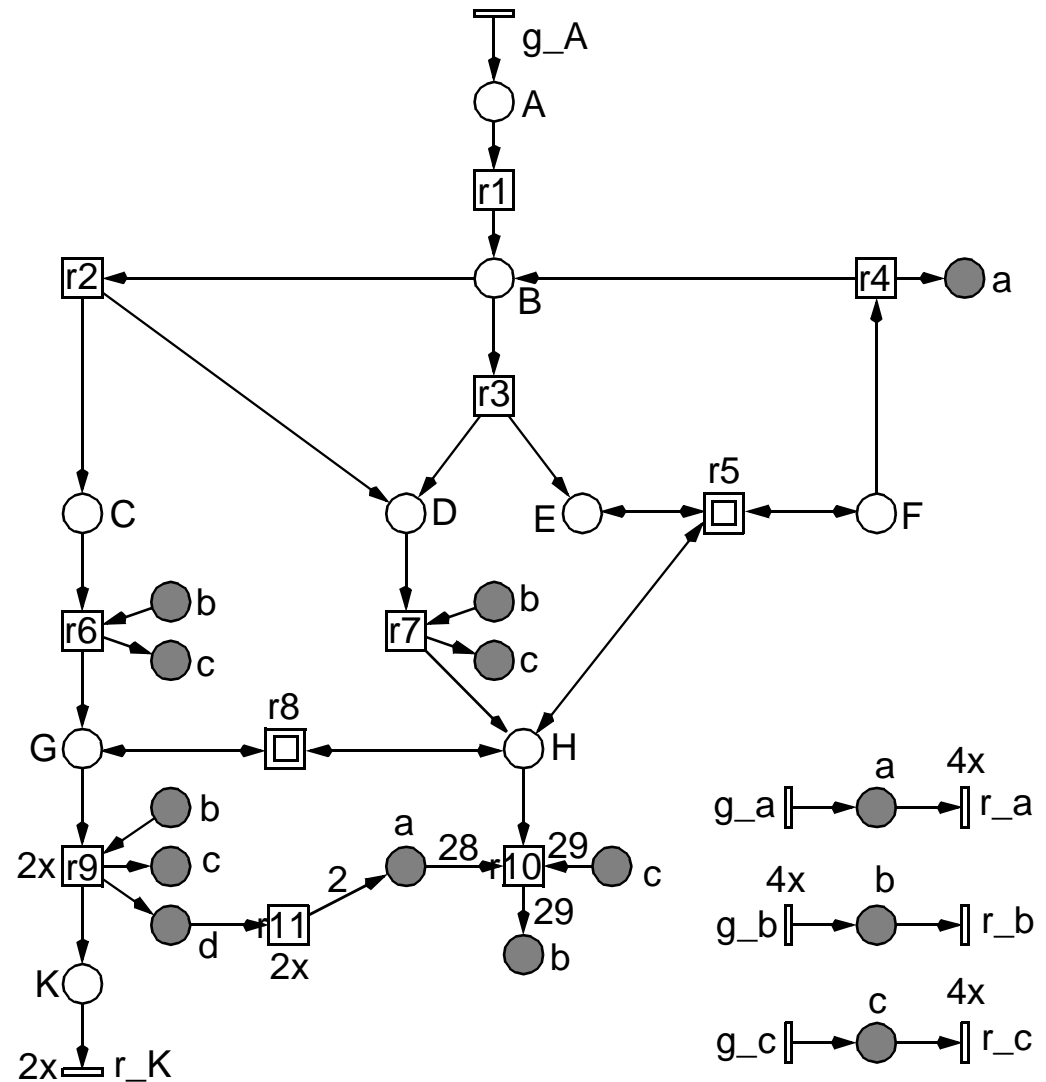


□ i/o-T-invariant, example

12		0.r1	:	1
		1.r2	:	1,
		3.r8_rev	:	1,
		4.r6	:	1,
		5.r7	:	1,
		9.r9	:	2,
		12.r11	:	2,
		13.g_A	:	1,
		14.r_K	:	2,
		15.g_b	:	4,
		18.r_c	:	4,
		20.r_a	:	4

□ sum equation

$$A + 4b \rightarrow 2K + 4a + 4c$$

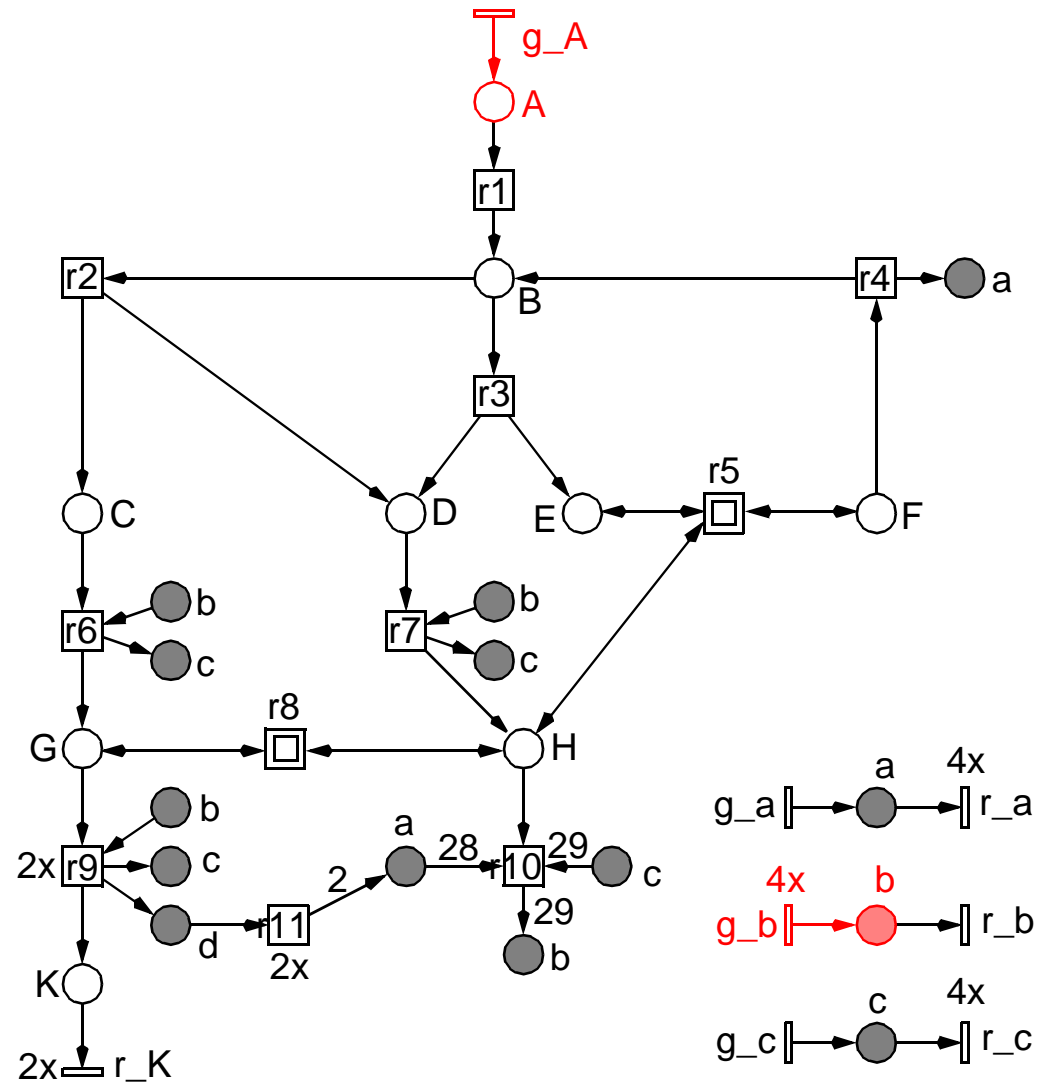


□ i/o-T-invariant, example

12		0.r1	:	1
		1.r2	:	1,
		3.r8_rev	:	1,
		4.r6	:	1,
		5.r7	:	1,
		9.r9	:	2,
		12.r11	:	2,
		13.g_A	:	1,
		14.r_K	:	2,
		15.g_b	:	4,
		18.r_c	:	4,
		20.r_a	:	4

□ sum equation

$$A + 4b \rightarrow 2K + 4a + 4c$$

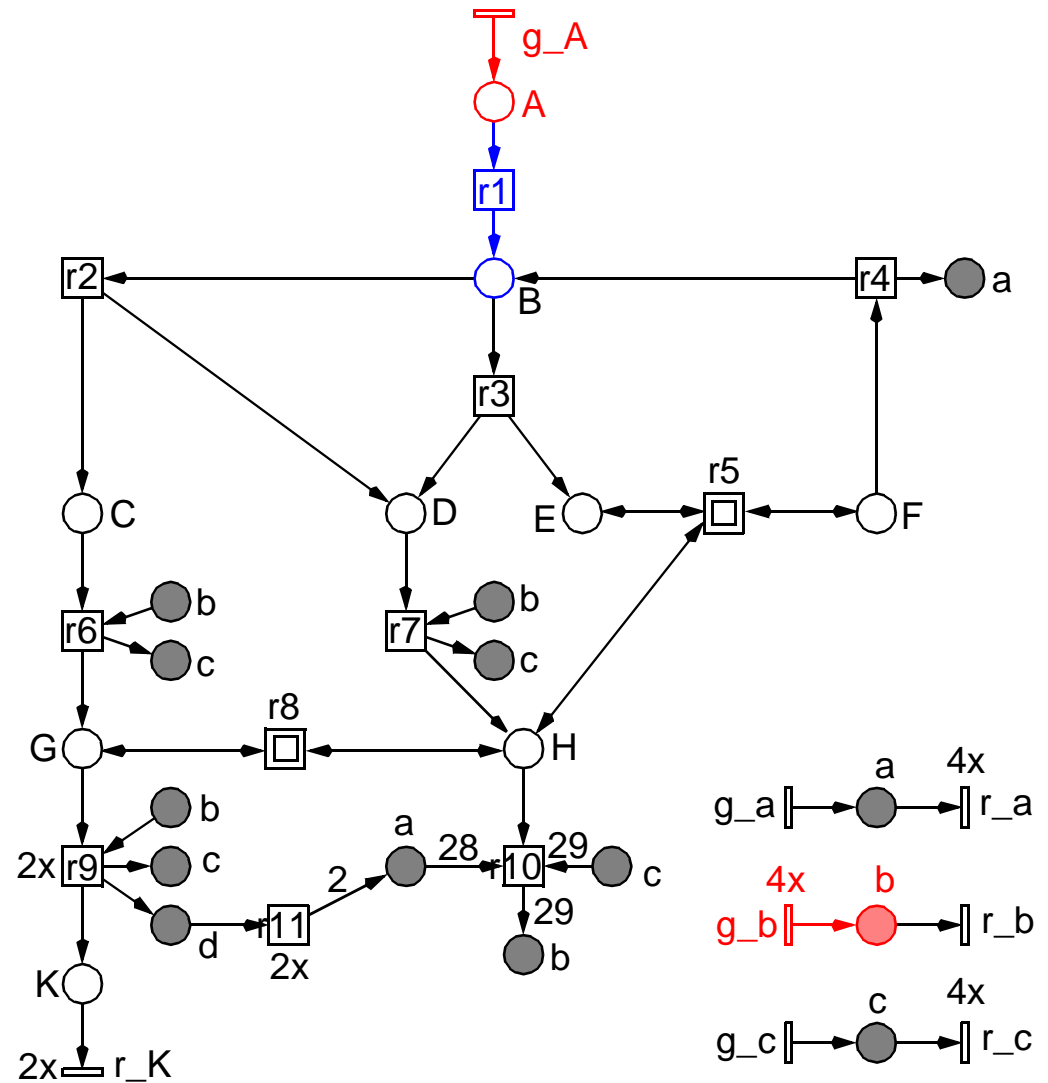


□ i/o-T-invariant, example

12		0.r1	:	1
		1.r2	:	1,
		3.r8_rev	:	1,
		4.r6	:	1,
		5.r7	:	1,
		9.r9	:	2,
		12.r11	:	2,
		13.g_A	:	1,
		14.r_K	:	2,
		15.g_b	:	4,
		18.r_c	:	4,
		20.r_a	:	4

□ sum equation

$$A + 4b \rightarrow 2K + 4a + 4c$$

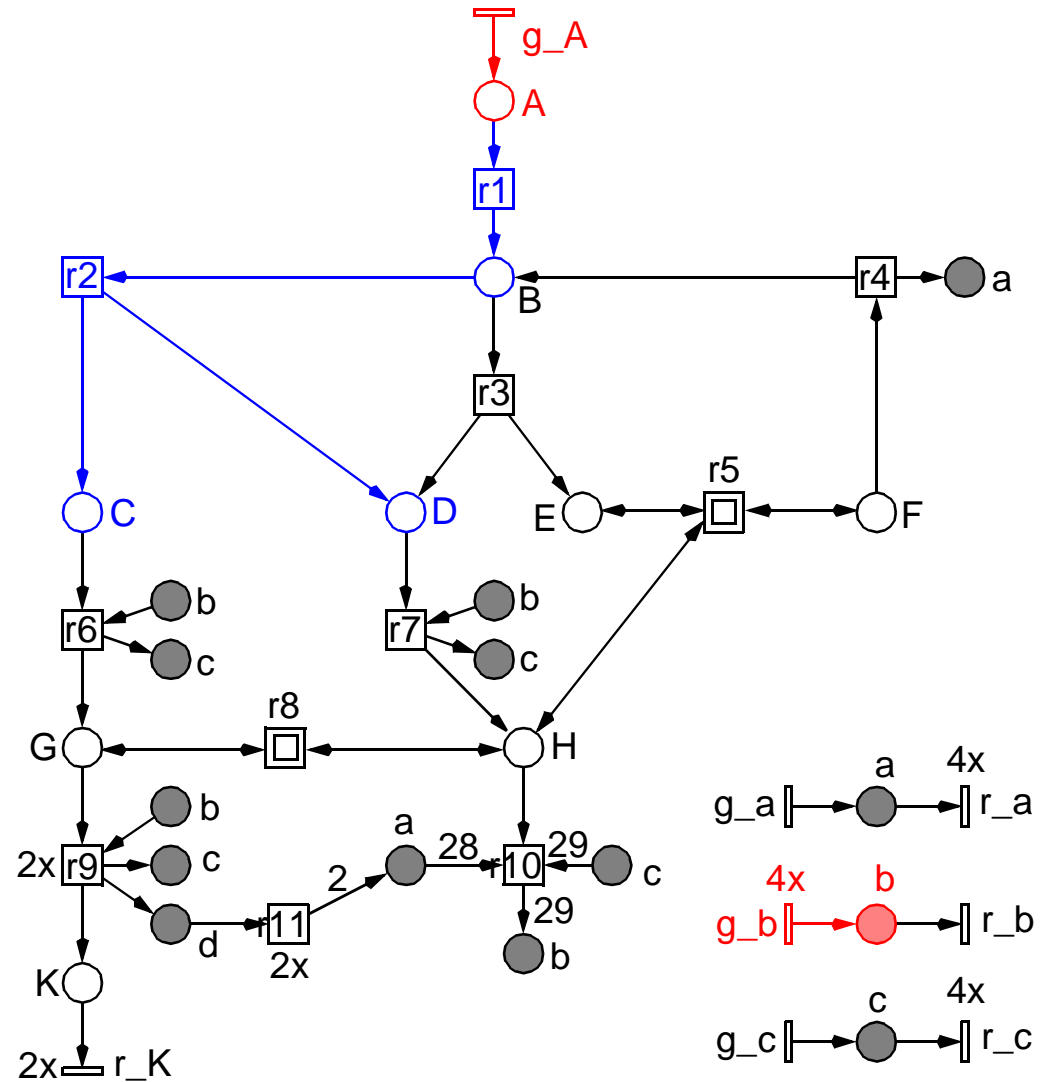


□ i/o-T-invariant, example

12		0.r1	:	1
		1.r2	:	1,
		3.r8_rev	:	1,
		4.r6	:	1,
		5.r7	:	1,
		9.r9	:	2,
		12.r11	:	2,
		13.g_A	:	1,
		14.r_K	:	2,
		15.g_b	:	4,
		18.r_c	:	4,
		20.r_a	:	4

□ sum equation

$$A + 4b \rightarrow 2K + 4a + 4c$$

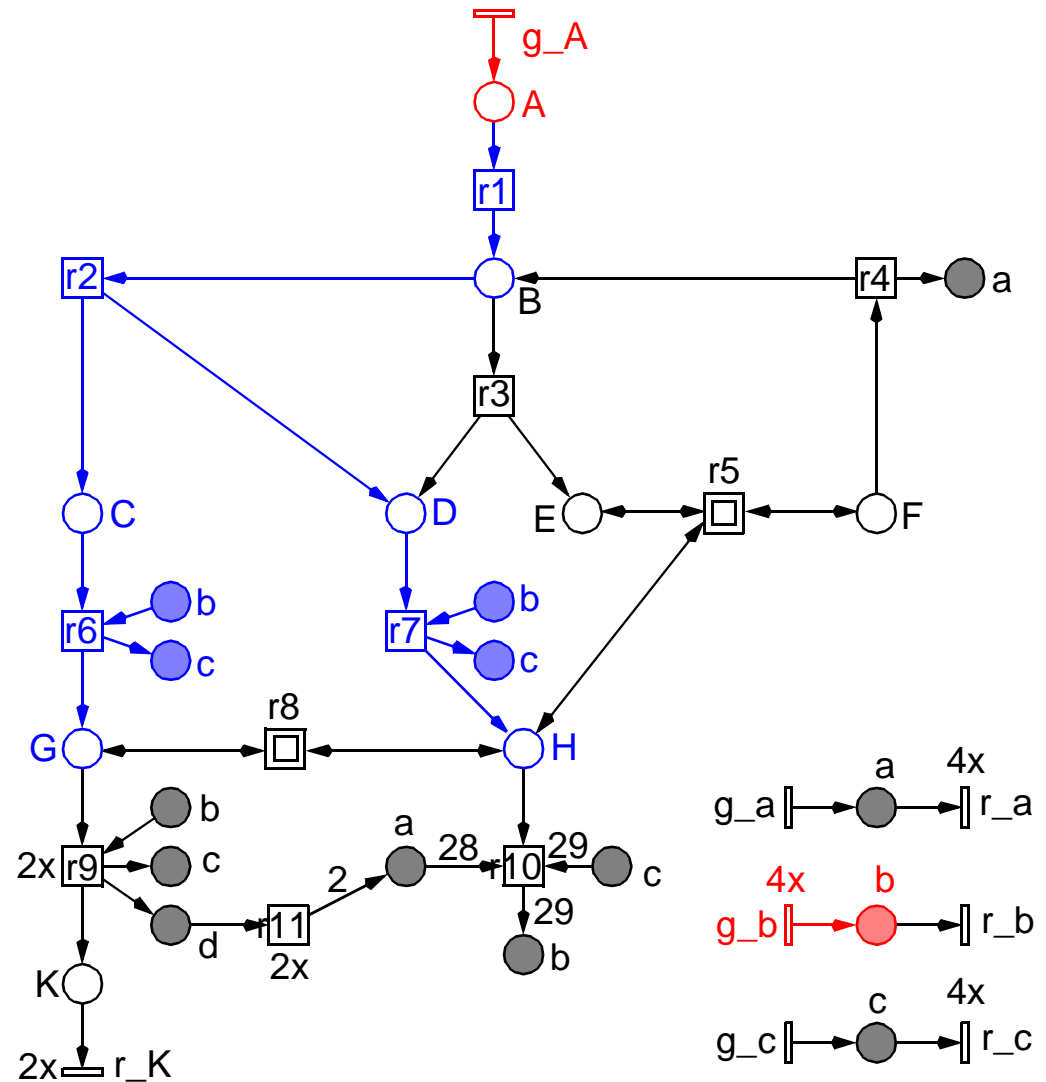


□ i/o-T-invariant, example

12		0.r1	:	1
		1.r2	:	1,
		3.r8_rev	:	1,
		4.r6	:	1,
		5.r7	:	1,
		9.r9	:	2,
		12.r11	:	2,
		13.g_A	:	1,
		14.r_K	:	2,
		15.g_b	:	4,
		18.r_c	:	4,
		20.r_a	:	4

□ sum equation

$$A + 4b \rightarrow 2K + 4a + 4c$$

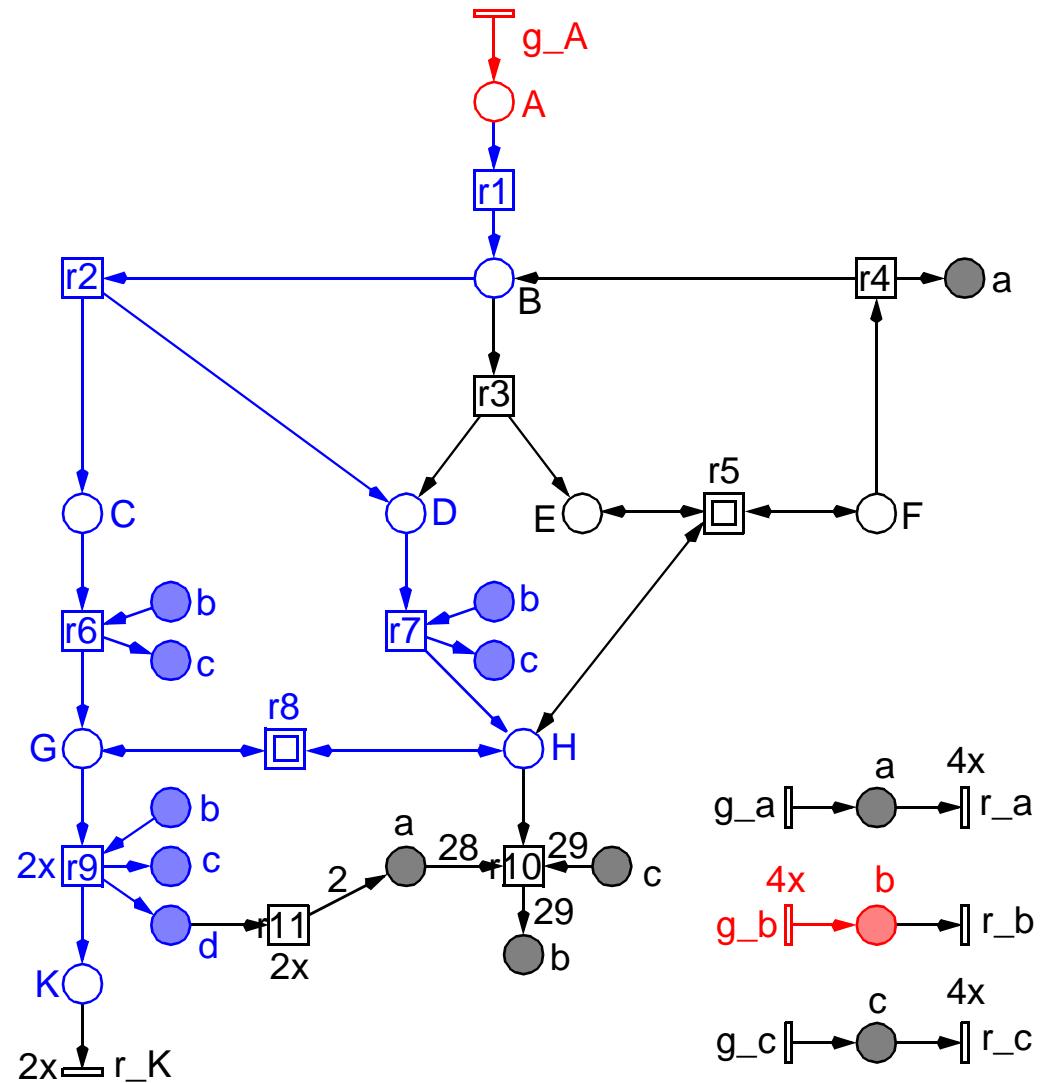


□ i/o-T-invariant, example

12		0.r1	:	1
		1.r2	:	1,
		3.r8_rev	:	1,
		4.r6	:	1,
		5.r7	:	1,
		9.r9	:	2,
		12.r11	:	2,
		13.g_A	:	1,
		14.r_K	:	2,
		15.g_b	:	4,
		18.r_c	:	4,
		20.r_a	:	4

□ sum equation

$$A + 4b \rightarrow 2K + 4a + 4c$$

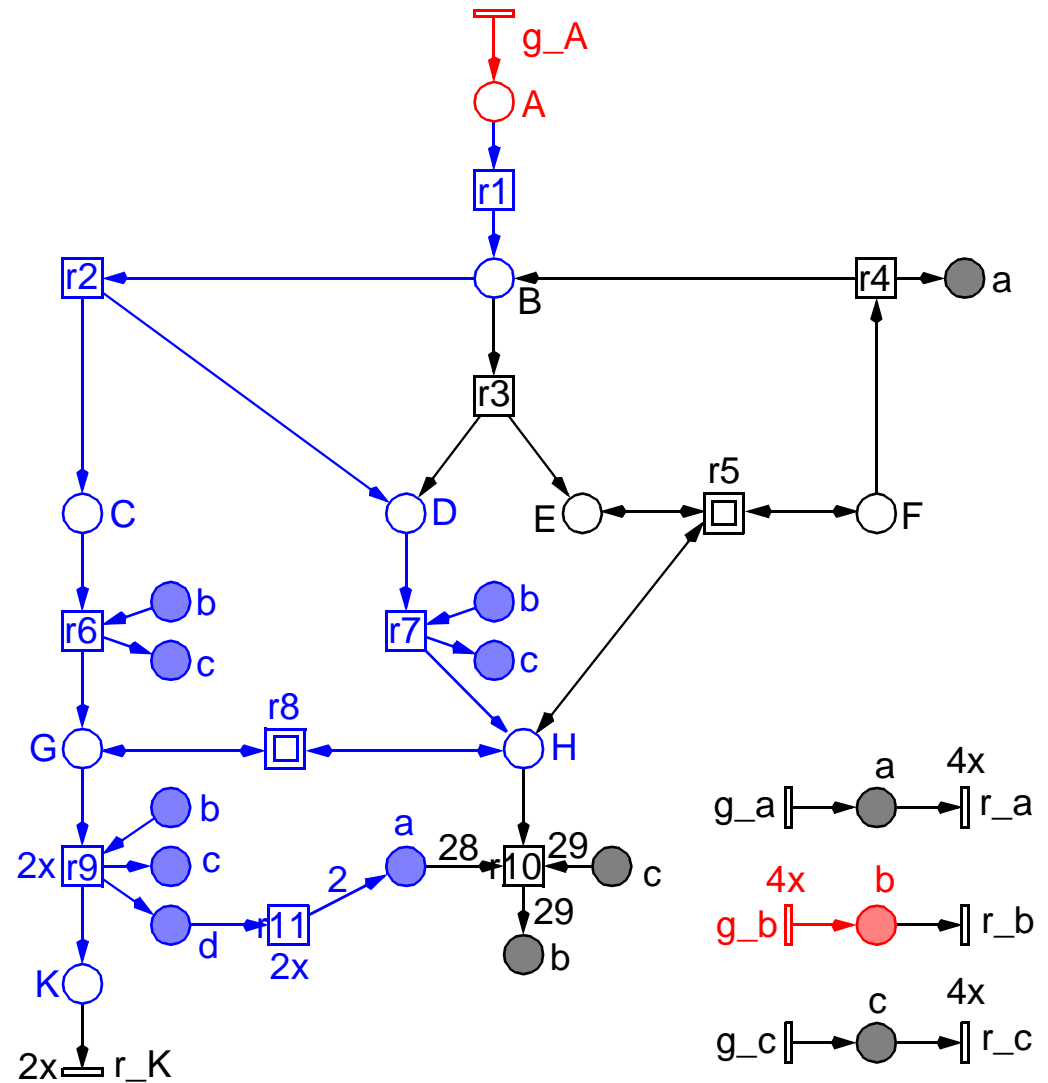


□ i/o-T-invariant, example

12		0.r1	:	1
		1.r2	:	1,
		3.r8_rev	:	1,
		4.r6	:	1,
		5.r7	:	1,
		9.r9	:	2,
		12.r11	:	2,
		13.g_A	:	1,
		14.r_K	:	2,
		15.g_b	:	4,
		18.r_c	:	4,
		20.r_a	:	4

□ sum equation

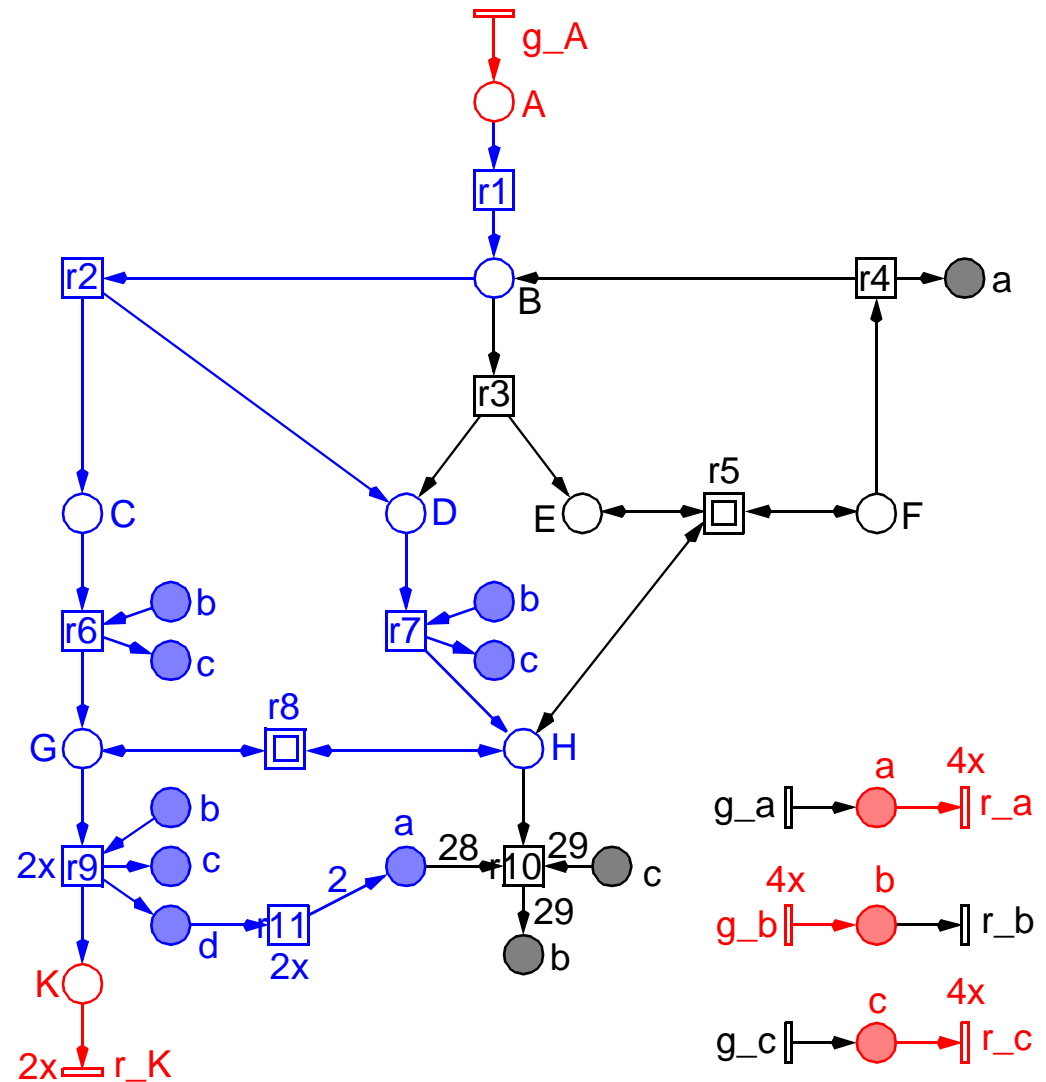
$$A + 4b \rightarrow 2K + 4a + 4c$$



□ i/o-T-invariant, example

12		0.r1	:	1
		1.r2	:	1,
		3.r8_rev	:	1,
		4.r6	:	1,
		5.r7	:	1,
		9.r9	:	2,
		12.r11	:	2,
		13.g_A	:	1,
		14.r_K	:	2,
		15.g_b	:	4,
		18.r_c	:	4,
		20.r_a	:	4

□ sum equation



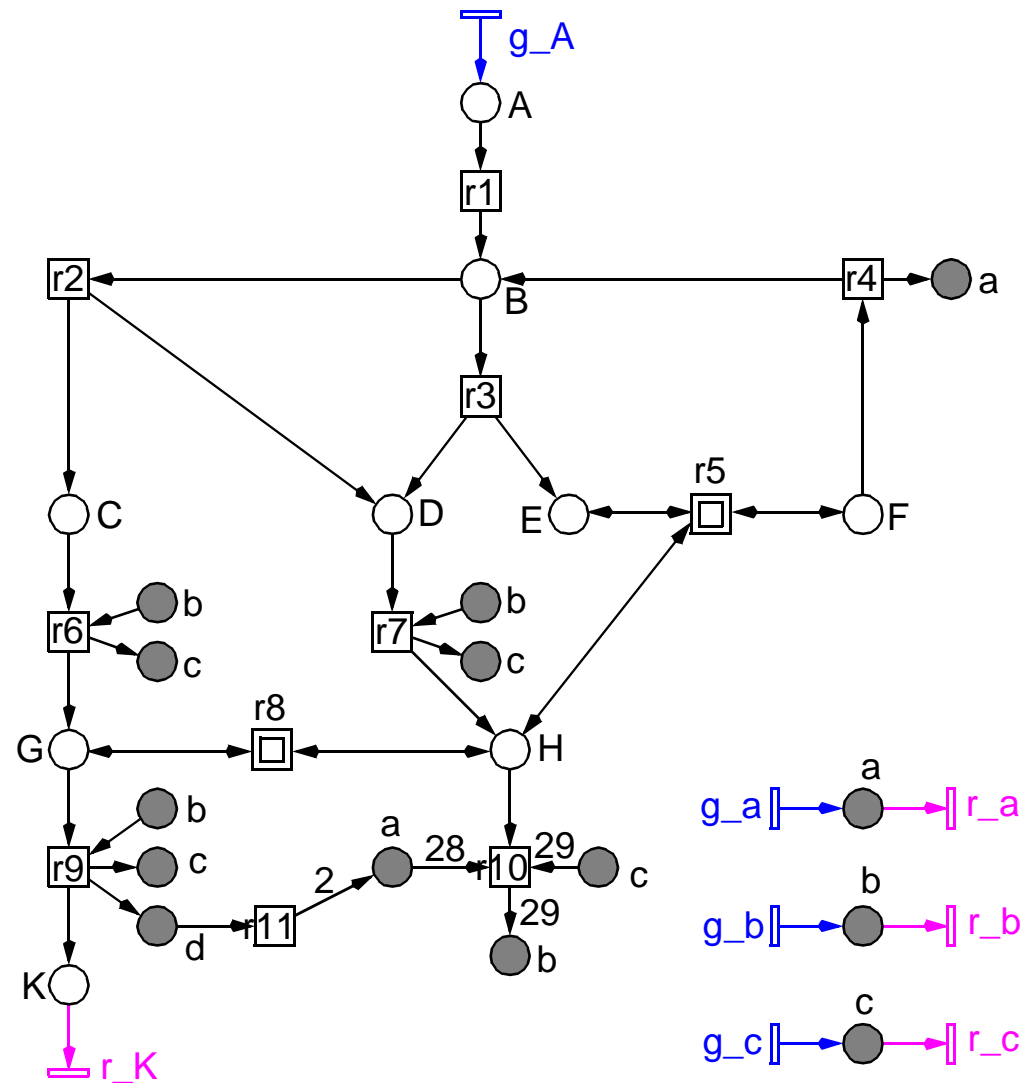
□ **validation criterion 1**

-> *CTI*

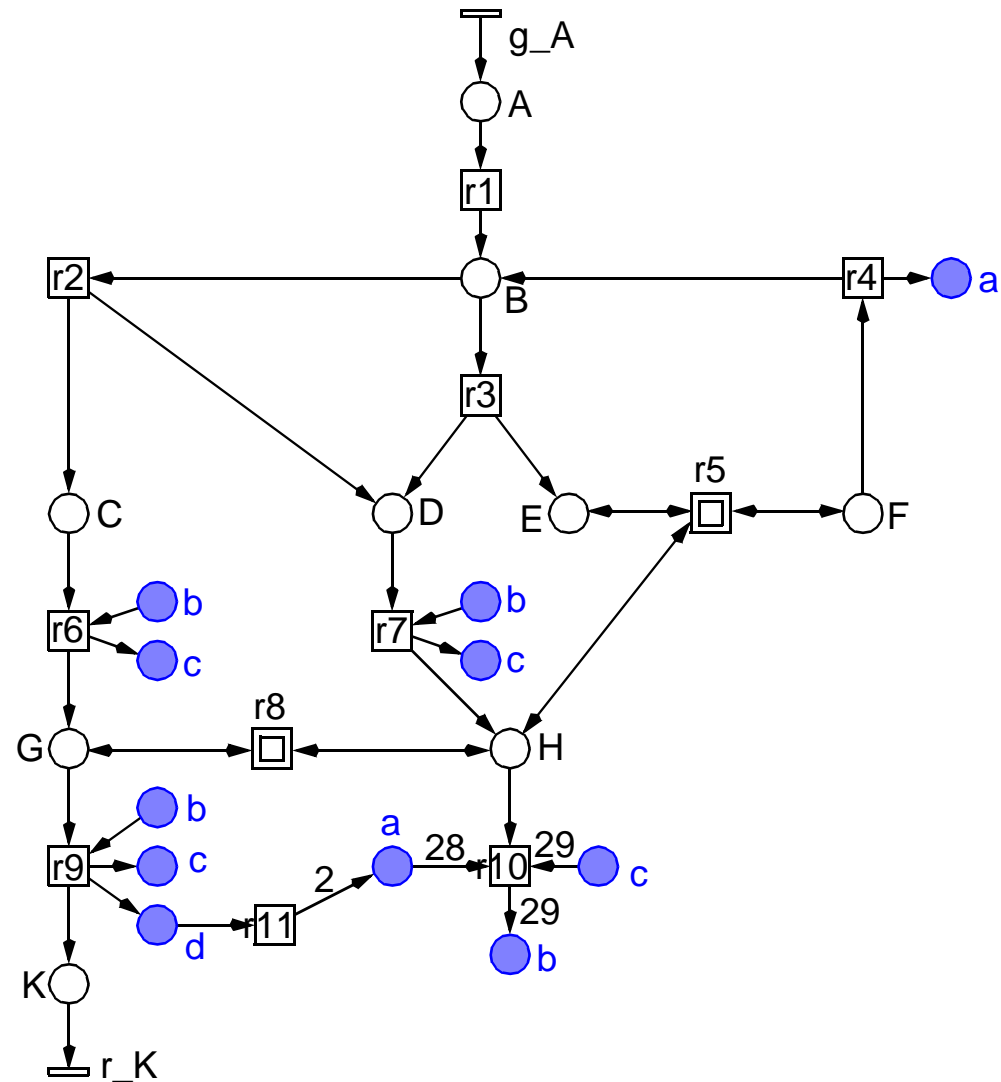
-> *no minimal T-invariant without biological interpretation*

-> *no known biological behaviour without corresponding T-invariant*

- ❑ **firm -> weak assumptions**
- ❑ **input substances**
-> *generating pre-transitions*
- ❑ **output substances**
-> *consuming post-transitions*



- ❑ **firm assumptions**
- ❑ **input substances**
-> *generating pre-transitions*
- ❑ **output substances**
-> *consuming post-transitions*
- ❑ **auxiliary substances**
-> *finite reservoir*
-> *P-invariants*
(= traps = co-traps)



- Lautenbach, 1973

- P-invariants

-> integer solutions of $yC = 0, y \neq 0, y \geq 0$

-> sets of places with

$$my = m_0y, \forall m \in RG(m_0)$$

- minimal P-invariants

-> there is no P-invariant with a smaller support

-> gcd of all entries is 1

↓
token conservation

- any P-invariant is a non-negative linear combination of minimal ones

-> multiplication with a positive integer

-> addition

-> Division by gcd

- Covered by P-Invariants (CPI)

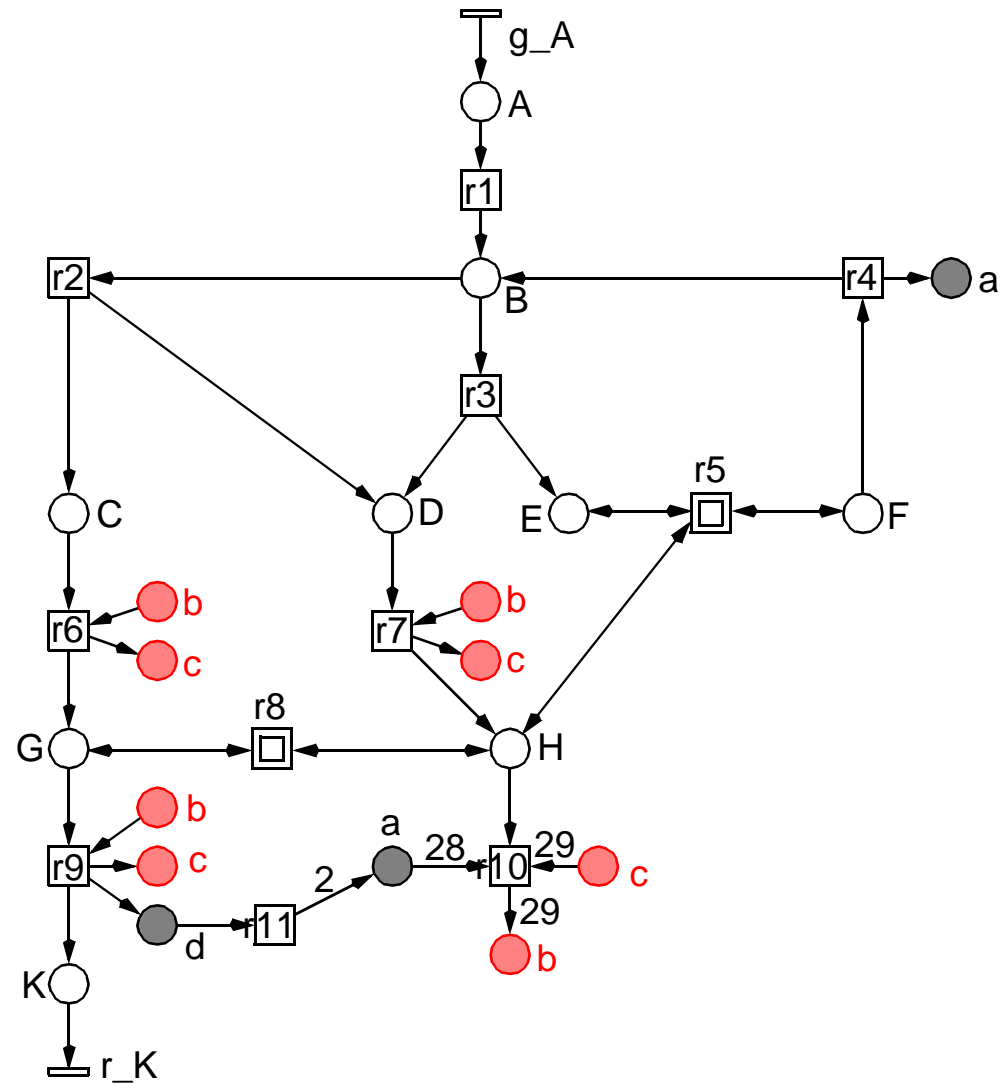
-> each place belongs to a P-invariant

□ P-invariants

-> *substance conservation*

□ P-invariants, examples

1 | 7.b : 1,
 | 8.c : 1



□ P-invariants

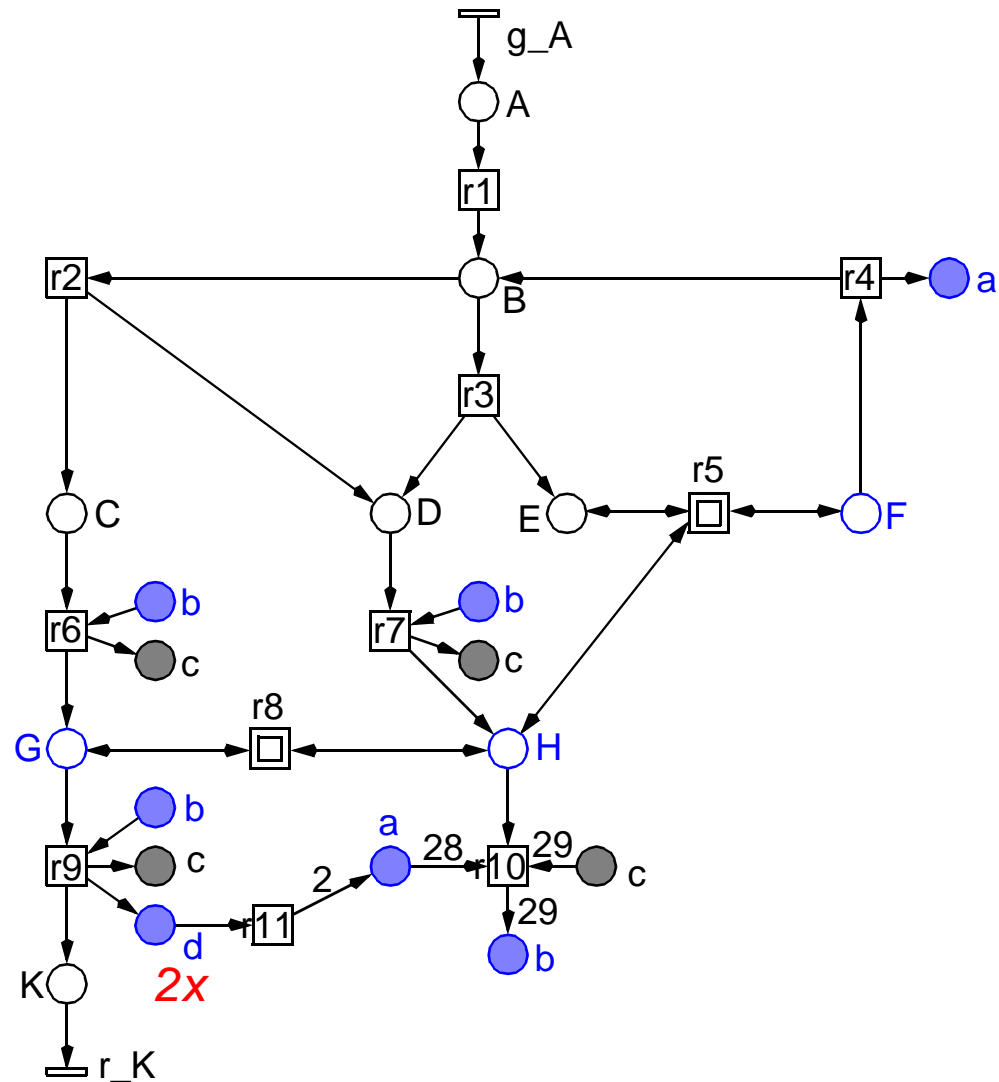
-> *substance conservation*

□ P-invariants, examples

1 | 7.b : 1,
 | 8.c : 1

2 | 5.G : 1,
 | 6.H : 1,
 | 7.b : 1,
 | 9.F : 1,
 | 10.a : 1,
 | 11.d : 2

□ Which (minimal) initial token distribution makes the net live ?



❑ validation criterion 1

-> *CTI*

-> *no minimal T-invariant without biological interpretation*

-> *no known biological behaviour without corresponding T-invariant*

❑ validation criterion 2

-> *P-invariants - groups of compounds with conservation property*

-> *no minimal P-invariant without biological interpretation*

❑ strong assumptions

- > quantitative relations of input/output compounds
- > finite reservoir of auxiliary compounds

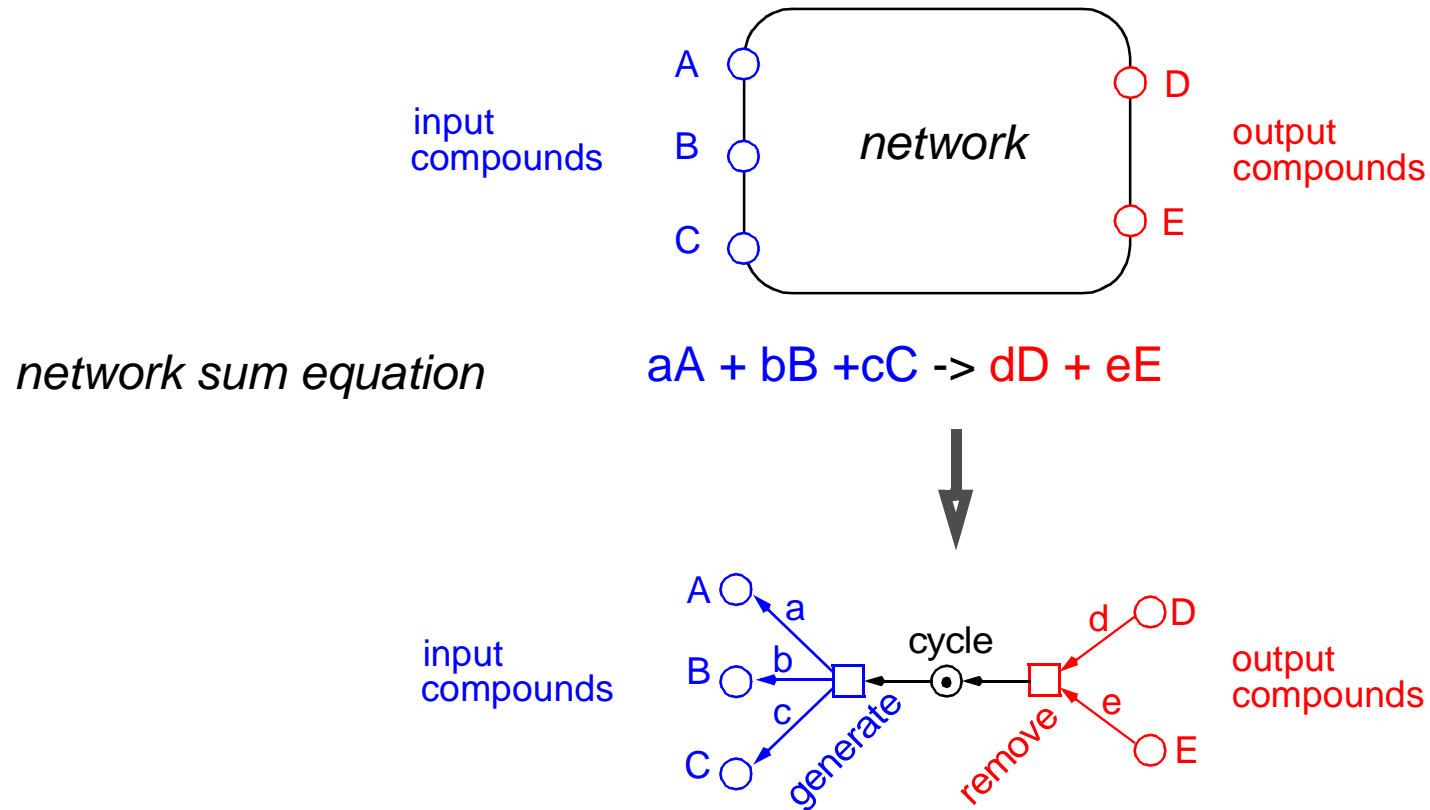
❑ typically expected properties

INA																	
ORD	HOM	NBM	PUR	CSV	SCF	CON	SC	Ft0	tF0	Fp0	pF0	MG	SM	FC	EFC	ES	
N	N	N	Y	N	N	Y	Y	N	N	N	N	N	N	N	N	N	
DTP	CPI	CTI	B	SB	REV	DSt	BSt	DTr	DCF	L	LV	L&S					
?	Y	Y	Y	Y	?	N	?	N	Y	Y	Y	N					

❑ boundedness (finite state space)

- > model checking of specific properties
 - CTL (undecidable for unbounded nets)
 - LTL (decidable, but no tools, not yet ?)
- > preserving all possible behavior (= minimal T-invariants)

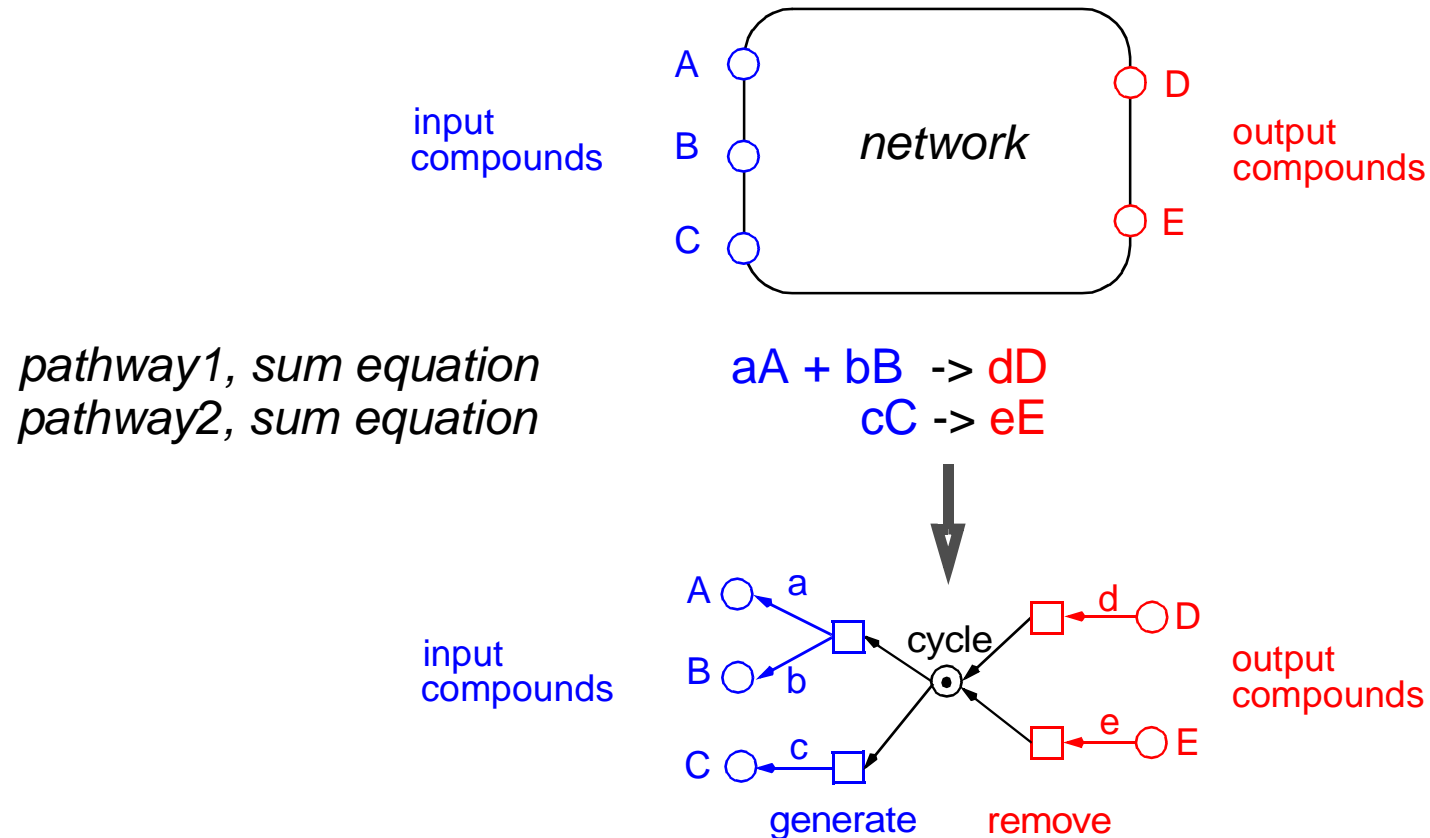
□ additional model component



□ precondition

-> equal sum equation for all pathways

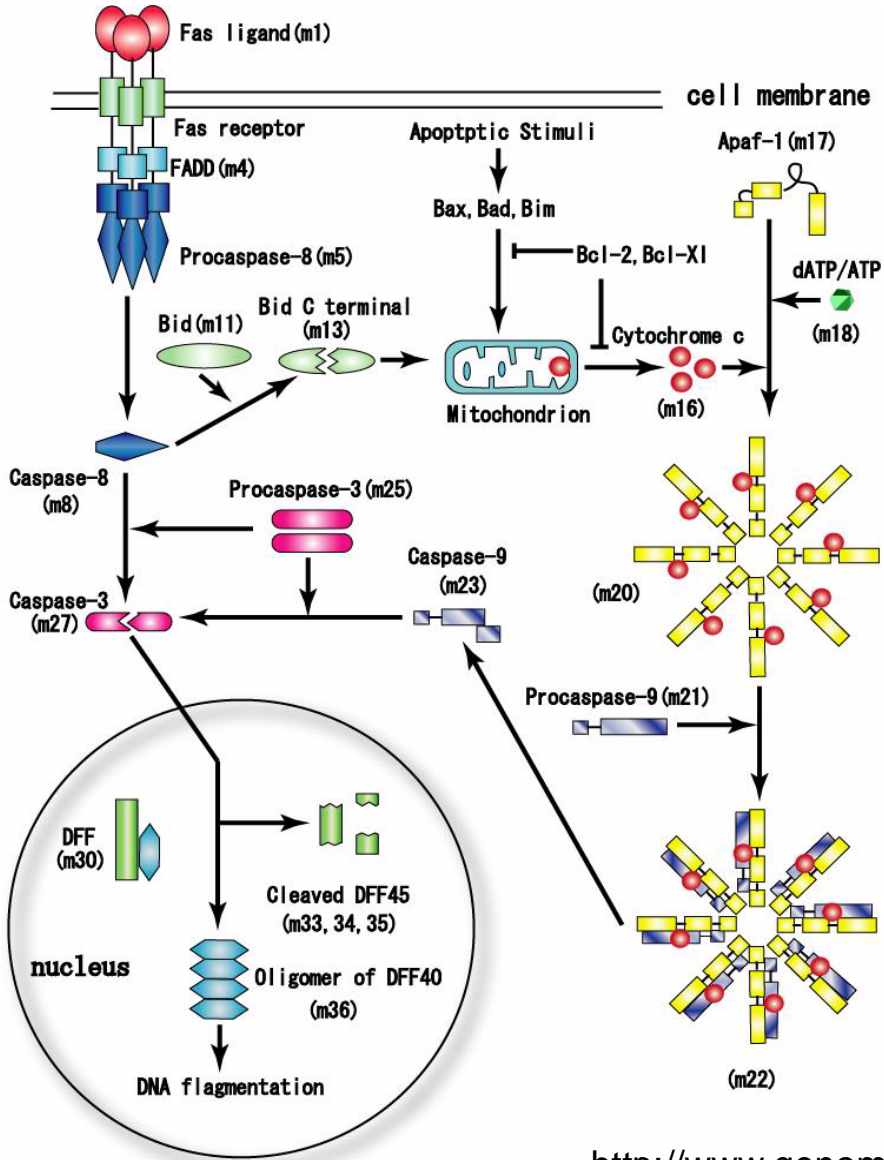
□ additional model component, refinement



□ **precondition**

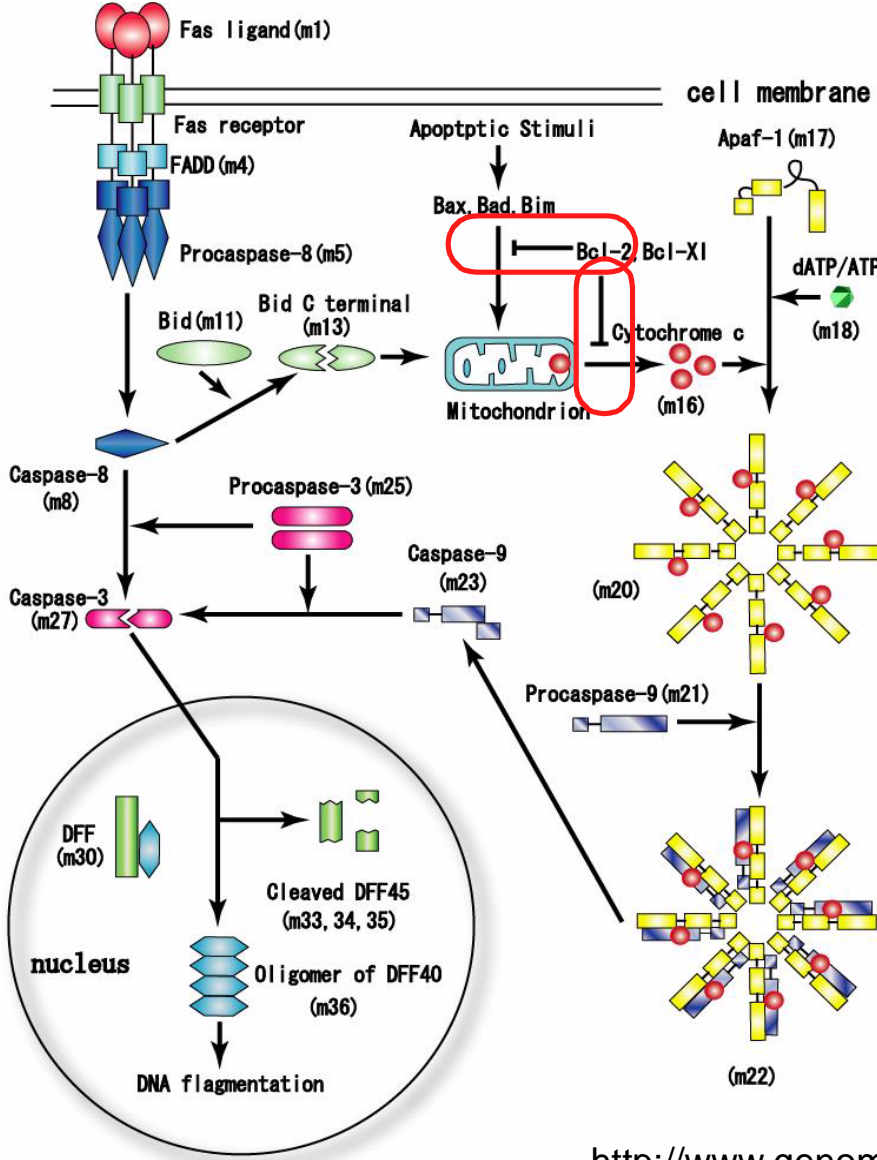
-> *controlled conflicts between pathways with unequal sum equations*

- example - apoptosis
-> Matsuno et al.
- signal-transduction pathway



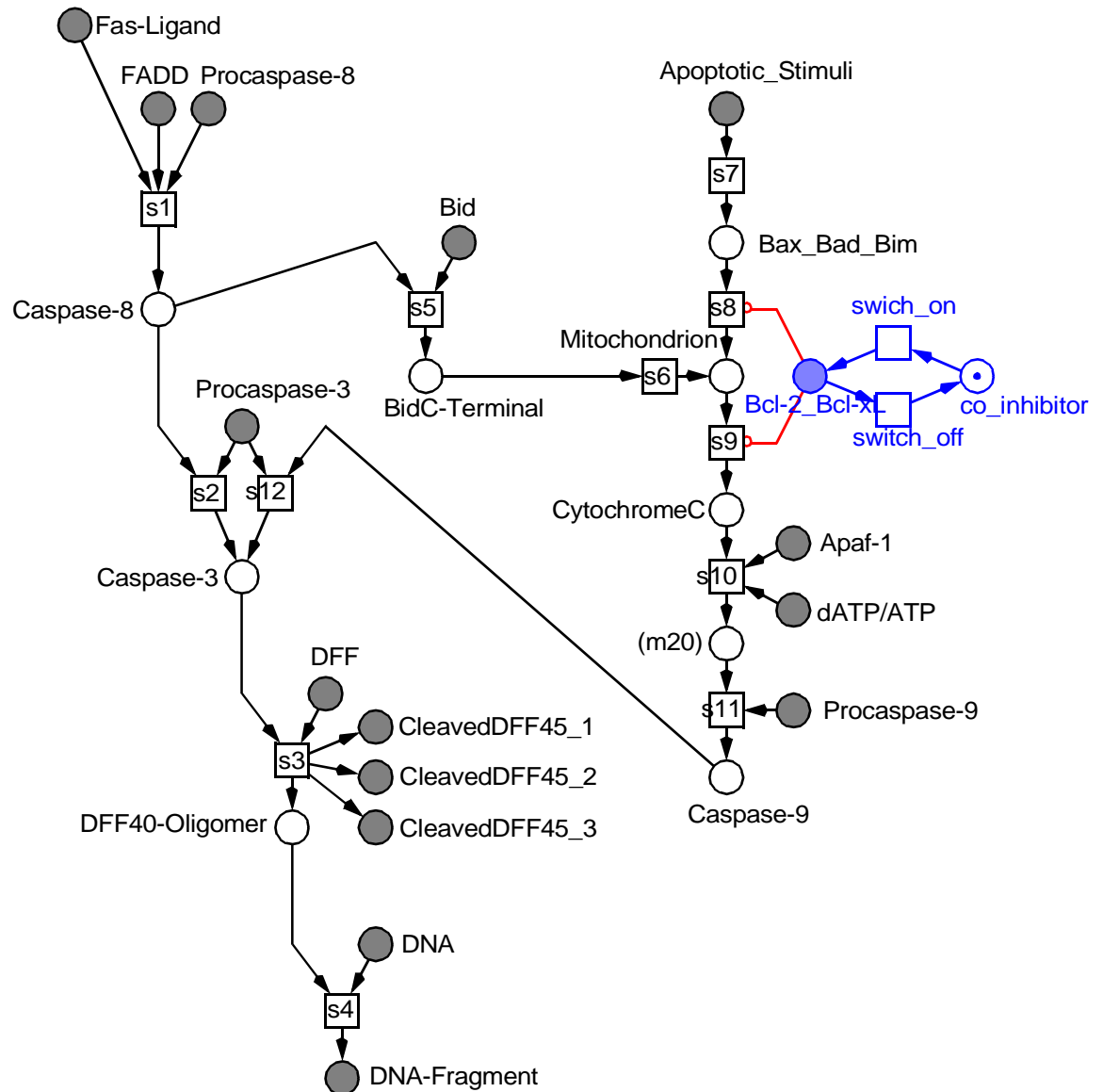
<http://www.genomicObject.net>

- example - apoptosis
-> Matsuno et al.
- signal-transduction pathway
- inhibitor arcs

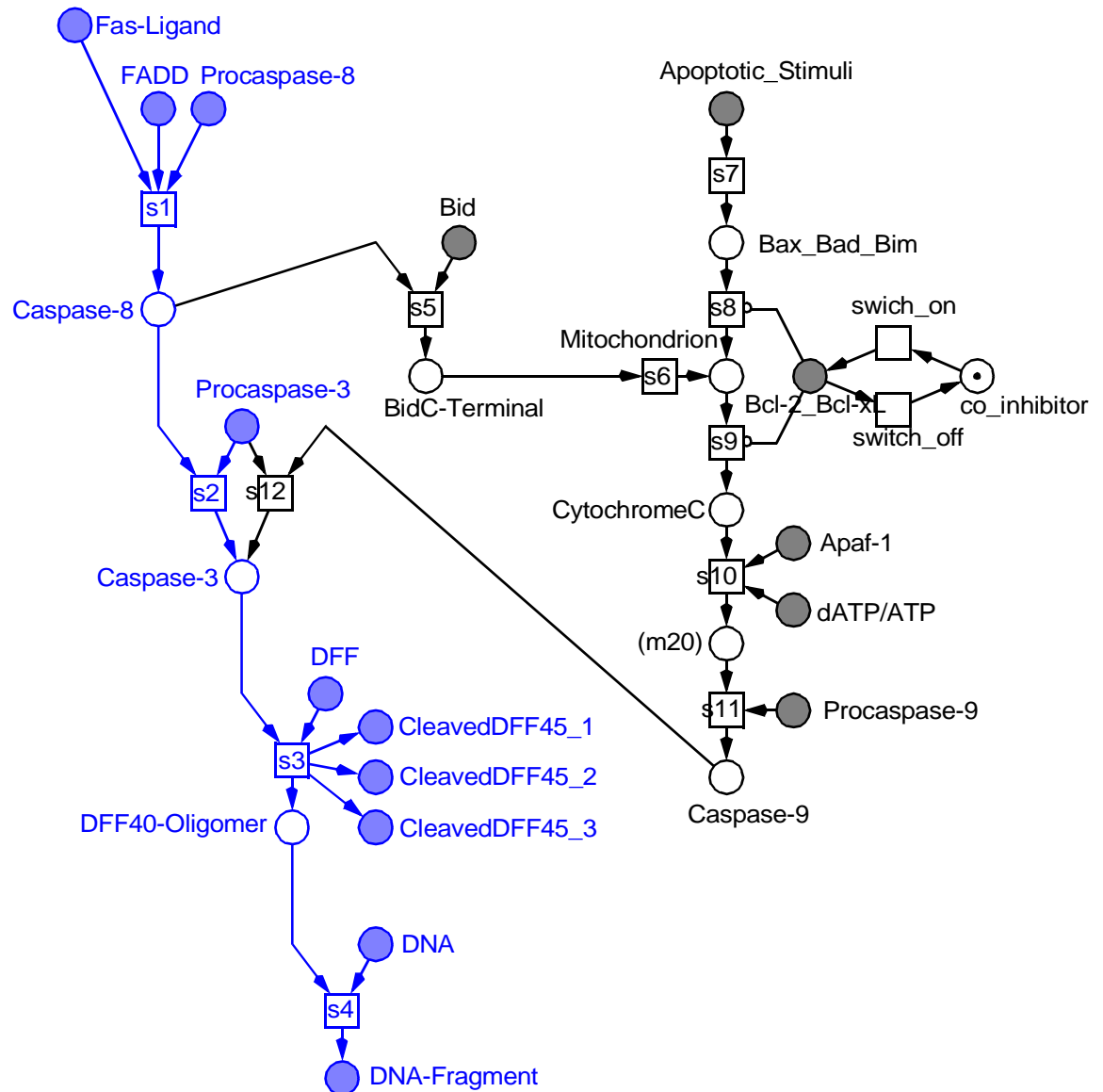


<http://www.genomicObject.net>

- example - apoptosis
- network model
- inhibitor arcs
- three pathways = min. T-invariants

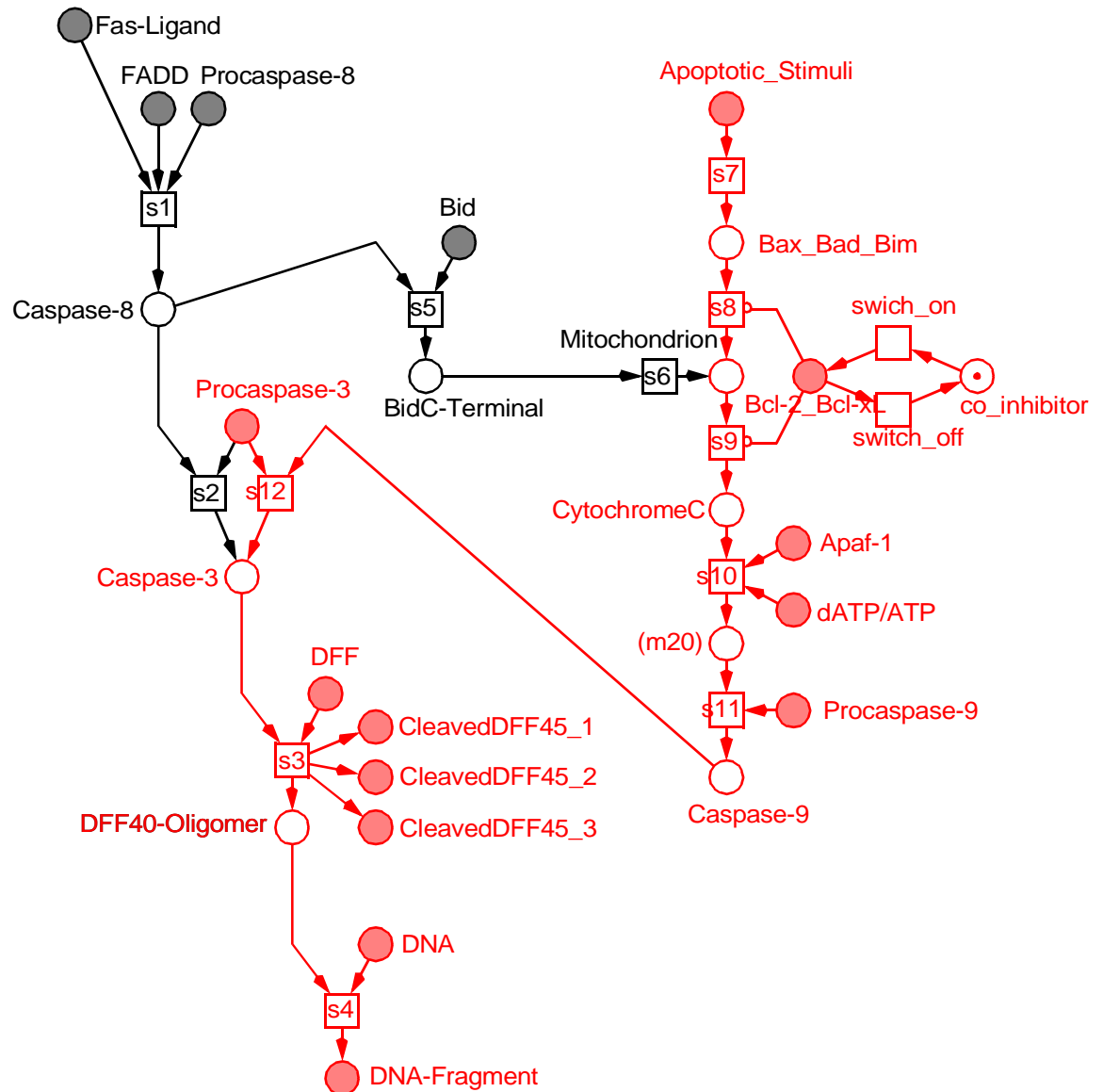


- example - apoptosis
- network model
- environment, style 1
 - > three pathways
 - = min. T-invariants
- T-invariant 1
 - > Fas-induced
 - > 'death-receptor' pathway

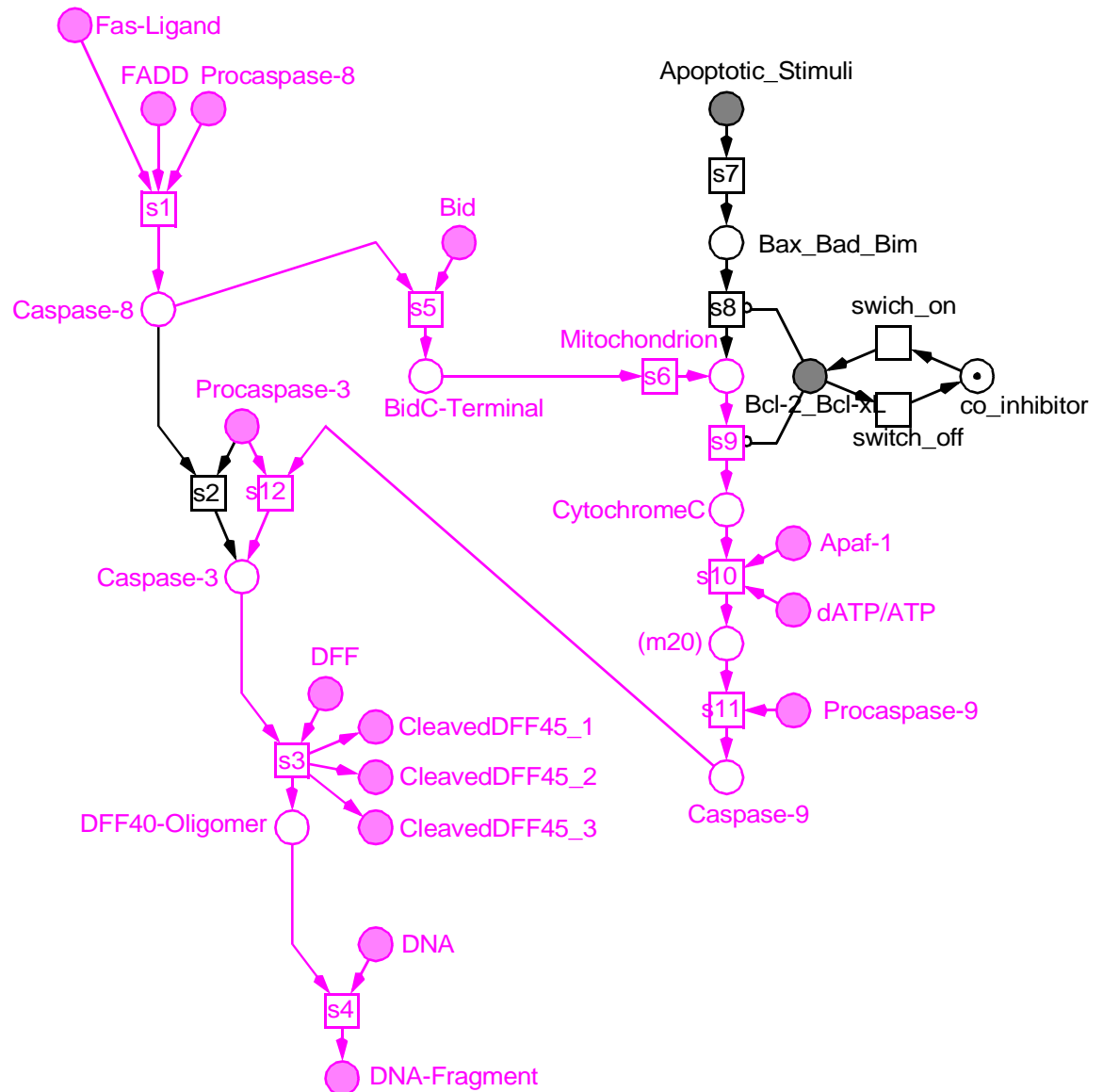


- example - apoptosis
- network model
- environment, style 1
-> three pathways
= min. T-invariants

- **T-invariant 2**
-> apoptotic-stimuli-induced
-> 'mitochondrial' pathway



- example - apoptosis
- network model
- environment, style 1
 - > three pathways
 - = min. T-invariants
- T-invariant 3
 - > 'cross-talk by Bid' pathway



□ example - apoptosis

□ environment model

□ pathway 1 / 3

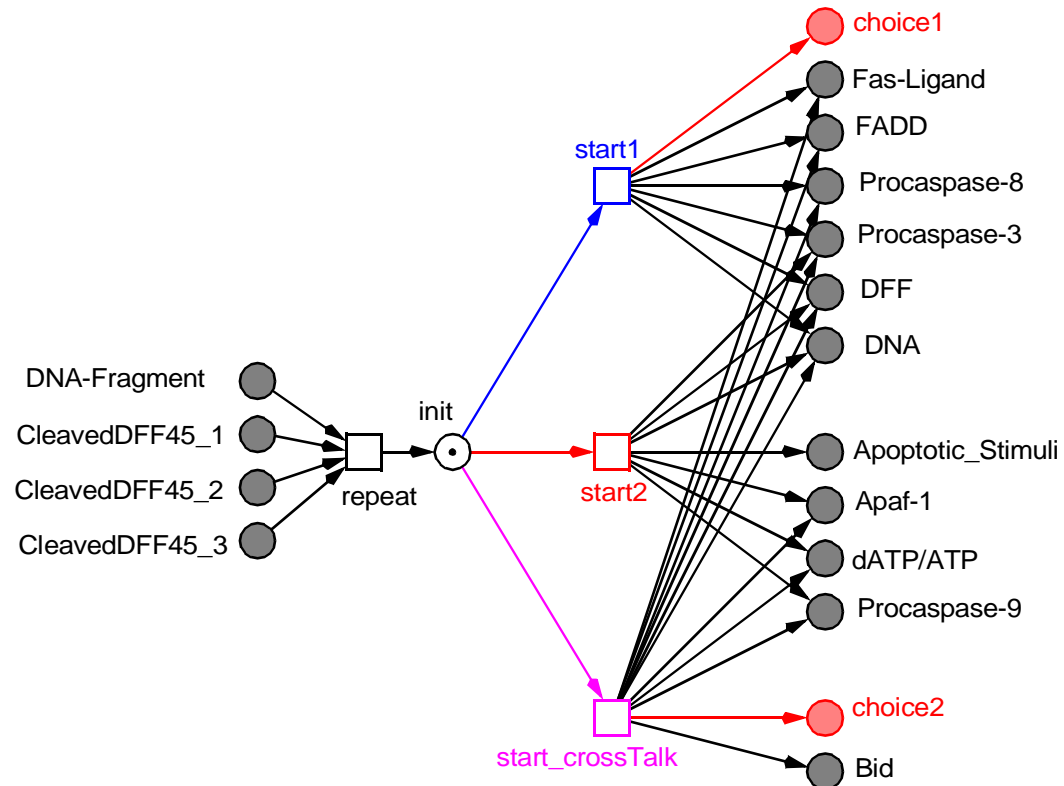
-> overlap at the beginning

-> then branch

-> *controlled by places choice1 / choice2*

□ all pathways share the same ending

-> *only one repeat transition*



□ example - apoptosis

□ network model,
adapted

□ system model

-> network model

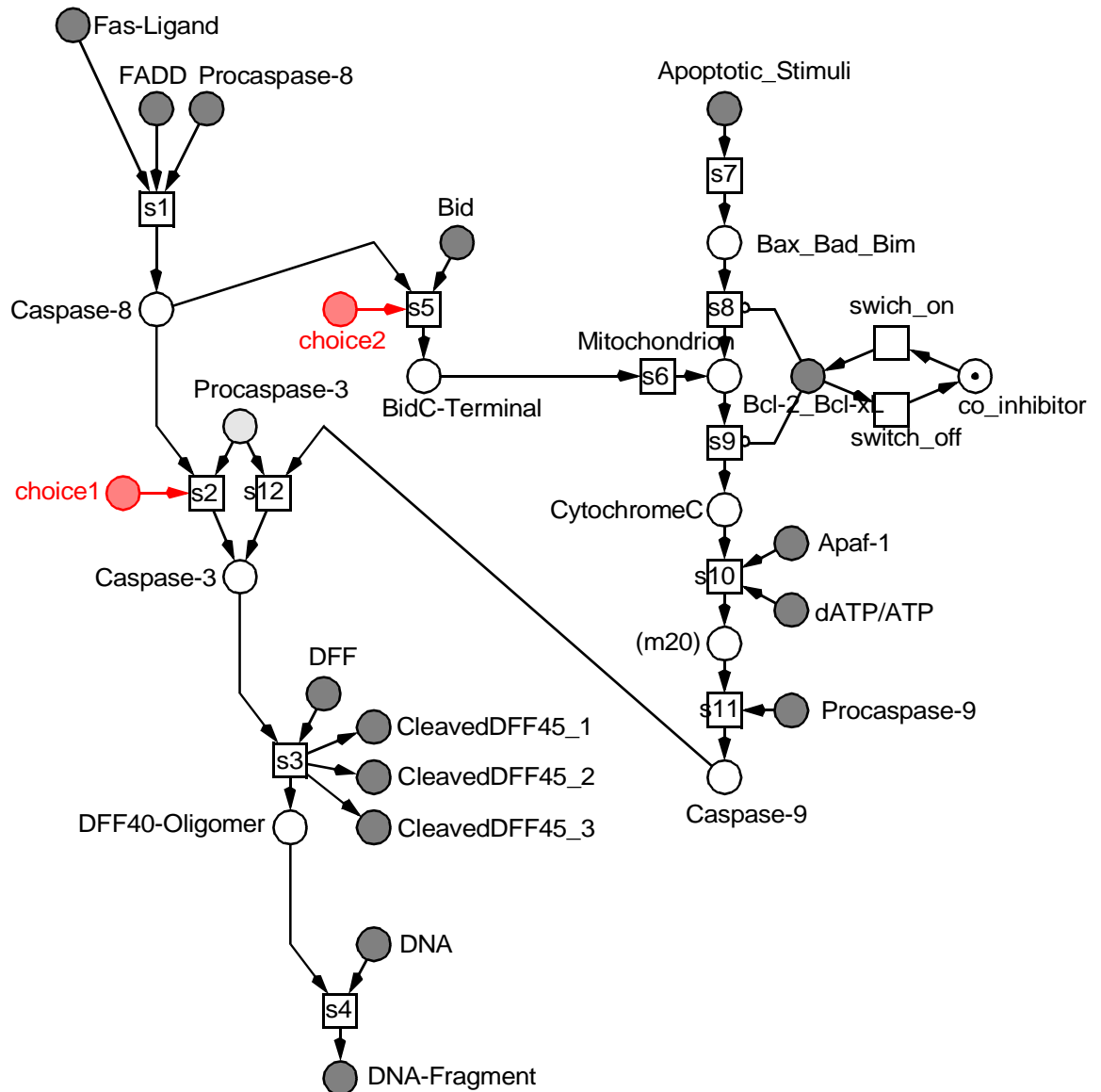
-> environment model

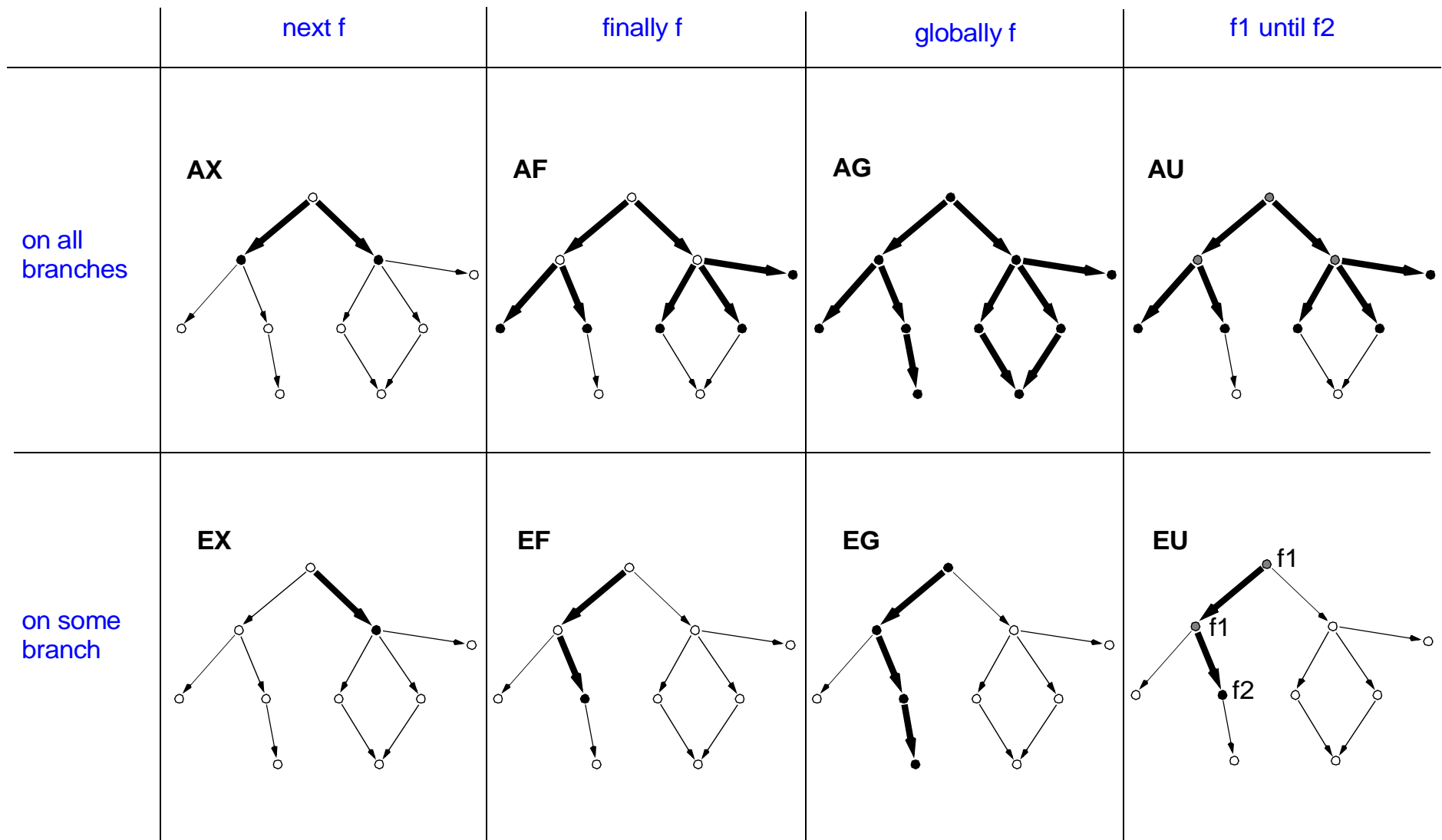
□ system model

-> 1-bounded

-> live

□ ready for
model checking





❑ property 1

if inhibitor substance $Bcl_2_Bcl_xL$ is present,
then the progress of the cross-talk pathway is stopped at *Mitochondrion*

$AG (Bcl_2_Bcl_xL * Mitochondrion \rightarrow AX (Mitochondrion));$

❑ property 2

if inhibitor substance $Bcl_2_Bcl_xL$ is present,
then the progress of the cross-talk pathway is stopped at *Mitochondrion*
until the inhibitor substance disappears

$AG (Bcl_2_Bcl_xL * Mitochondrion \rightarrow A (Mitochondrion \ U \ ! \ Bcl_2_Bcl_xL));$

❑ property 3

if inhibitor substance $Bcl_2_Bcl_xL$ is not present,
then the progress of the cross-talk pathway is not stopped at *Mitochondrion*

$AG (!Bcl_2_Bcl_xL * Mitochondrion \rightarrow EX (CytochromeC));$

❑ **validation criterion 1**

-> *CTI*

-> *no minimal T-invariant without biological interpretation*

-> *no known biological behaviour without corresponding T-invariant*

❑ **validation criterion 2**

-> *P-invariants - groups of compounds with conservation property*

-> *no minimal P-invariant without biological interpretation*

❑ **validation criterion 3**

-> *CPI*

-> *all expected temporal-logic properties -> TRUE*

- ❑ **extensions**
 - > *read arcs*
 - > *inhibitor arcs !?*

- ❑ **efficient computation of minimal invariants**
 - > *exponential complexity*
 - > *compositional / step-wise refinement approach ?*

- ❑ **analysis of bounded, but not safe non-ordinary nets with inhibitor arcs**
 - > *huge state spaces, beyond exponential growth (?)*
 - > *smaller, bounded version of case study 2 $\geq 10^{10}$ states (IDD-based mc tool)*

- ❑ **analysis of unbounded nets**
 - > *besides T-invariant analysis ?*

- ❑ **model checking**
 - > *relevant properties ?*

- ❑ **representation of bionetworks by Petri nets**
 - > *unifying view*
 - > *animation*
 - > *model validation against consistency criteria*
 - > *qualitative/quantitative behaviour prediction*

- ❑ **qualitative model -> steady state behavior**

- ❑ **three styles of environment description**
 - > *style 1* -> *T-invariants <-> pathways, CTI*
 - > *style 2* -> *P-invariants <-> substance conservation*
 - > *style 3* -> *CPI, model checking of CTL - properties*

- ❑ **many challenging questions for analysis techniques**

THANKS !