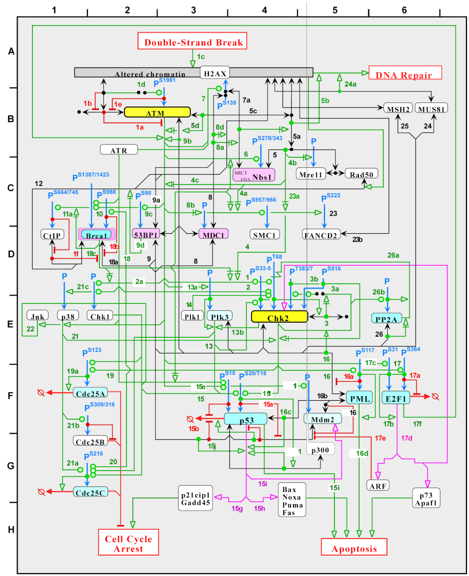
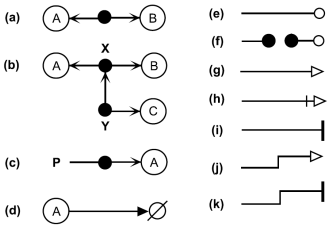
# Molecular Interaction Automated Maps

Robert Demolombe
Luis Fariñas del Cerro
Naji Obeid

Université de Toulouse and CNRS, IRIT, Toulouse, France

# Defining the Problem




Metabolic network formed by long sequences of positive (activation) and negative (inhibition) biochemical reactions.

# Content

## Content

- Logical model capable of describing general metabolic pathways and their possible extensions.
- Translation procedure for eliminating first order variables and equality predicates.

# Content

## Content

- Logical model capable of describing general metabolic pathways and their possible extensions.
- Translation procedure for eliminating first order variables and equality predicates.

# Language

## State Predicates

- $A(x)$: $x$ is *Active*.
- $I(x)$: $x$ is *Inhibited*.
- $P(x)$: $x$ is *Present*.

# Language

## State Predicates

- $A(x)$: $x$ is *Active*.
- $I(x)$: $x$ is *Inhibited*.
- $P(x)$: $x$ is *Present*.

## State Axioms

$$\neg(A(x) \wedge I(x)) \ . \tag{1}$$

$$P(x) \leftrightarrow A(x) \vee I(x) \ . \tag{2}$$

# Language - Action Predicates

## Capacity of Activation

$CA(y, x)$: $y$ can activate $x$.

# Language - Action Predicates

## Capacity of Activation

$CA(y, x)$: $y$ can activate $x$.

## Effective Capacity of Activation

$CA^e(y, x)$: $y$ can effectively activate $x$.

# Language - Action Predicates

## Capacity of Activation

$CA(y, x)$: $y$ can activate $x$.

## Effective Capacity of Activation

$CA^e(y, x)$: $y$ can effectively activate $x$.

## Direct or Indirect Capacity of Activation

$CA^{di}(y, x)$: $y$ can directly or indirectly activate $x$.

# Language - Action Predicates

### Capacity of Activation

$CA(y, x)$: $y$ can activate $x$.

### Effective Capacity of Activation

$CA^e(y, x)$: $y$ can effectively activate $x$.

### Direct or Indirect Capacity of Activation

$CA^{di}(y, x)$: $y$ can directly or indirectly activate $x$.

### Capacity to Inhibit the Capacity of Activation

$CICA(z, y, x)$: $z$ can inhibit the capacity that $y$ has to activate $x$.

# Language - Action Predicates

## Capacity of Inhibition

$CI(y, x)$: $y$ can inhibit $x$.

# Language - Action Predicates

## Capacity of Inhibition

$CI(y, x)$: $y$ can inhibit $x$.

## Effective Capacity of Inhibition

$CI^e(y, x)$: $y$ can effectively inhibit $x$.

# Language - Action Predicates

### Capacity of Inhibition

$CI(y, x)$: $y$ can inhibit $x$.

### Effective Capacity of Inhibition

$CI^e(y, x)$: $y$ can effectively inhibit $x$.

### Direct or Indirect Capacity of Inhibition

$CI^{di}(y, x)$: $y$ can directly or indirectly inhibit $x$.

# Language - Action Predicates

## Capacity of Inhibition

$CI(y,x)$: $y$ can inhibit $x$.

## Effective Capacity of Inhibition

$CI^e(y,x)$: $y$ can effectively inhibit $x$.

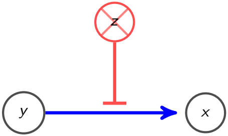
## Direct or Indirect Capacity of Inhibition

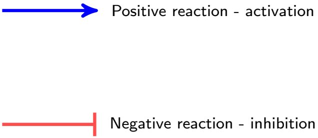
$CI^{di}(y,x)$: $y$ can directly or indirectly inhibit $x$.

## Capacity to Inhibit the Capacity of Inhibition

$CICI(z,y,x)$: $z$ can inhibit the capacity that $y$ has to inhibit $x$.

# Language - Activation Axiom



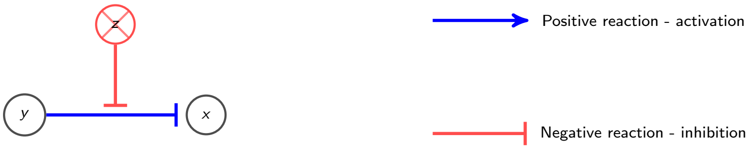

## Activation Axiom

$$\forall x \forall y (A(y) \wedge CA^e(y,x) \rightarrow A(x)) \quad (3)$$

With:

$$CA^e(y,x) \stackrel{\text{def}}{=} CA(y,x) \wedge \neg \exists z (CICA(z,y,x) \wedge A(z)) \quad (4)$$

# Language - Inhibition Axiom




## Inhibition Axiom

$$\forall x \forall y (A(y) \wedge CI^e(y,x) \rightarrow I(x)) \quad (5)$$

With:

$$CI^e(y,x) \stackrel{\text{def}}{=} CI(y,x) \wedge \neg\exists z (CICI(z,y,x) \wedge A(z)) \quad (6)$$

# Language - Causal Relations - Activation

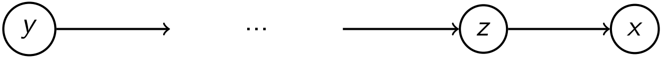



Figure : Direct or Indirect Capacity of Activation

## Activation

From

$$\forall x \forall y (CA^{e}(y,z) \vee \exists z (CA^{di}(y,z) \wedge CA^{e}(z,x)) \leftrightarrow CA^{di}(y,x)). \qquad (7)$$

We can deduce:

$$\forall x \forall y (A(y) \wedge CA^{di}(y,x) \rightarrow A(x)) \qquad (8)$$

# Language - Causal Relations - Inhibition

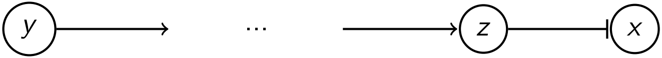



Figure : Direct or Indirect Capacity of Inhibition

## Inhibition

From

$$\forall x \forall y (CI^{e}(y,z) \vee \exists z (CA^{di}(y,z) \wedge CI^{e}(z,x)) \leftrightarrow CI^{di}(y,x)). \quad (9)$$

We can deduce:

$$\forall x \forall y (A(y) \wedge CI^{di}(y,x) \rightarrow I(x)) \quad (10)$$

# Language - Extension - Phosphorylation

## Capacity of Phosphorylation

$CP(z, y, s, x)$: $z$ can phosphorylate $y$ on site $s$, where $x$ is the result of the phosphorylation.

# Language - Extension - Phosphorylation

## Capacity of Phosphorylation

$CP(z, y, s, x)$: $z$ can phosphorylate $y$ on site $s$, where $x$ is the result of the phosphorylation.

## Effective Capacity of Phosphorylation

$CP^e(z, y, s, x)$: $z$ can effectively phosphorylate $y$ on site $s$, where $x$ is the result of the phosphorylation.

# Language - Extension - Phosphorylation

## Capacity of Phosphorylation

$CP(z, y, s, x)$: $z$ can phosphorylate $y$ on site $s$, where $x$ is the result of the phosphorylation.

## Effective Capacity of Phosphorylation

$CP^e(z, y, s, x)$: $z$ can effectively phosphorylate $y$ on site $s$, where $x$ is the result of the phosphorylation.

## Direct or Indirect Capacity of Phosphorylation

$CP^{di}(z, y, s, x)$: $z$ can directly or indirectly phosphorylate $y$ on site $s$, where $x$ is the result of the phosphorylation.

# Language - Extension - Phosphorylation

## Capacity of Phosphorylation

$CP(z, y, s, x)$: $z$ can phosphorylate $y$ on site $s$, where $x$ is the result of the phosphorylation.

## Effective Capacity of Phosphorylation

$CP^e(z, y, s, x)$: $z$ can effectively phosphorylate $y$ on site $s$, where $x$ is the result of the phosphorylation.

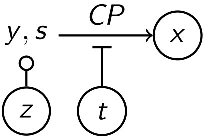
## Direct or Indirect Capacity of Phosphorylation

$CP^{di}(z, y, s, x)$: $z$ can directly or indirectly phosphorylate $y$ on site $s$, where $x$ is the result of the phosphorylation.

## Capacity to Inhibit the Capacity of Phosphorylation

$CICP(t, z, y, s, x)$: $t$ can inhibit the capacity that $z$ has to phosphorylate $y$.

# Language - Extension - Phosphorylation




## Activation Axiom

$$\forall x \forall y \forall s \forall z (A(z) \wedge A(y) \wedge CP^e(z, y, s, x) \rightarrow A(x)) \quad (11)$$

With:

$$CP^e(z, y, s, x) \stackrel{\text{def}}{=} CP(z, y, s, x) \wedge \neg \exists t (CICP(t, z, y, s, x) \wedge A(z)) \quad (12)$$

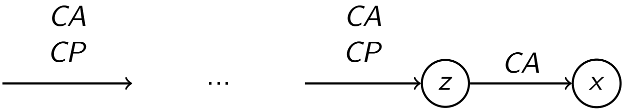



Figure : Direct or Indirect Capacity of Activation - Updated

## Activation

$$\forall x \forall y (CA^e(y,z) \vee \exists z (CA^{di}(y,z) \wedge CA^e(z,x)) \leftrightarrow CA^{di}(y,x)).$$

And

$$\forall x \forall y (CA^e(y,z) \vee \exists w \exists s \exists z (CP^{di}(y,w,s,z) \wedge CA^e(z,x)) \leftrightarrow CA^{di}(y,x)).$$

# Language - Causal Relations - Phosphorylation

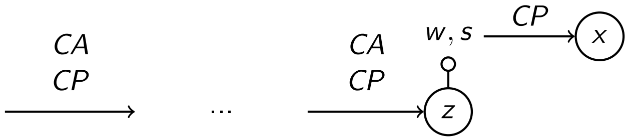



Figure : Direct or Indirect Capacity of Activation - Updated

## Activation

$$\forall x \forall y \forall w \forall s (CP^e(y, w, s, x) \vee \exists z (CA^{di}(y, z) \wedge CP^e(z, w, s, x)) \leftrightarrow CP^{di}(y, w, s, x))$$

And

$$\forall x \forall y \forall w \forall s (CP^e(y, w, s, x) \vee \exists w_1 \exists s_1 (CP^{di}(y, w_1, s_1, z) \wedge CP^e(z, w, s, x)) \leftrightarrow CP^{di}(y, w, s, x))$$

# Translation Procedure

## Domain formulas

$$\delta ::= P(\overline{x}, \overline{c}) | \varphi \vee \psi | \varphi \wedge \psi | \varphi \wedge \neg\psi \ . \quad (13)$$

Variables $\overline{x}$ and constants $\overline{c}$ denote $x_1, ..., x_n$ and $c_1, ..., c_m$ respectively.

The set of free variables in $\varphi$ is the same as the set of free variables in $\psi$ for $\varphi \vee \psi$.

The set of free variables in $\psi$ is included in the set of free variables in $\varphi$ for $\varphi \wedge \neg\psi$.

There are no special constraints for $\varphi \wedge \psi$.

# Translation Procedure

## Restricted formulas

$$\delta ::= \forall \overline{x}(\varphi \rightarrow \psi) | \exists \overline{x}(\varphi \wedge \psi) \ . \quad (14)$$

Where $\varphi$ is a domain formula and $\psi$ is either a restricted formula or a formula without quantifiers, and every variable appearing in a restricted formula must appear in a domain formula.

The set of variables in $\overline{x}$ is included in the set of free variables in $\varphi$; The same goes for $\psi$.

## Examples

$\forall x(P(x) \rightarrow Q(x))$.
$\forall x(P(x) \rightarrow \exists y(Q(y) \wedge R(x,y)))$.

# Translation Procedure

## Completion formulas

$$\forall x_1, ..., x_n \ (P(x_1, ..., x_n, c_1, ..., c_p) \leftrightarrow ((x_1 = a_{1_1} \wedge ... \wedge x_n = a_{1_n}) \vee ... \vee (x_1 = a_{m_1} \wedge ... \wedge x_n = a_{m_n}))) \ . \quad (15)$$

Where $P$ is a predicate symbol of arity $n + p$, and $a_i$ are constants.

## Definition

Given a domain formula $\varphi$ and a set of completion formulas $\alpha_1, ..., \alpha_n$ such that for each predicate symbol in $\varphi$ there exists a completion formula $\alpha$ for this predicate symbol, we say that the set of completion formulas $\alpha_1, ..., \alpha_n$ covers $\varphi$ and will be noted $C(\varphi)$.

# Translation Procedure

## Domain of the variables of a domain formula

- if $\varphi$ is of the form $P(x_1, ..., x_n, c_1, ..., c_p)$, and $C(\varphi)$ of the form:
$$\forall x_1, ..., x_m(P(x_1, ..., x_m, c_1, ..., c_l) \leftrightarrow ((x_1 = a_{1_1} \wedge ... \wedge x_m = a_{1_m}) \vee ... \vee (x_1 = a_{q_1} \wedge ... \wedge x_m = a_{q_m})))\ .$$
where $n \leq m$ and $l \leq p$.
then $$D(\mathcal{V}(\varphi), C(\varphi)) = \{< a_{1_1}, ..., a_{1_n} >, ..., < a_{q_1}, ..., a_{q_n} >\}\ . \quad (16)$$
- if $\varphi$ is of the form $\varphi_1 \vee \varphi_2$ then:
$$D(\mathcal{V}(\varphi_1 \vee \varphi_2), C(\varphi_1 \vee \varphi_2)) = D(\mathcal{V}(\varphi_1), C(\varphi_1)) \cup D(\mathcal{V}(\varphi_2), C(\varphi_2))\ . \quad (17)$$

# Translation Procedure

## Domain of the variables of a domain formula - Continued

- if $\varphi$ is of the form $\varphi_1 \wedge \varphi_2$ then:
$$D(\mathcal{V}(\varphi_1 \wedge \varphi_2), C(\varphi_1 \wedge \varphi_2)) = D(\mathcal{V}(\varphi_1), C(\varphi_1)) \otimes_c D(\mathcal{V}(\varphi_2), C(\varphi_2)) \quad . \tag{18}$$
Where $\otimes_c$ is a join operator and $c$ is a conjunction of equalities of the form $i = j$ where the same variable symbol appears in $\varphi_1 \wedge \varphi_2$ in position $i$ in $\varphi_1$ and in position $j$ in $\varphi_2$.

- if $\varphi$ is of the form $\varphi_1 \wedge \neg\varphi_2$ then:
$$D(\mathcal{V}(\varphi_1 \wedge \neg\varphi_2), C(\varphi_1 \wedge \neg\varphi_2)) = D(\mathcal{V}(\varphi_1), C(\varphi_1)) \setminus D(\mathcal{V}(\varphi_1 \wedge \varphi_2), C(\varphi_1 \wedge \varphi_2)). \tag{19}$$
Where $\setminus$ denotes the complement of the domain of each shared variable of $\varphi_1 \wedge \varphi_2$ with respect to $\varphi_1$.

# Translation Procedure

## Example

Considering the three domains formulas $P(x)$, $Q(x)$, $R(x,y)$ and their corresponding completion formulas as following:

$\forall x(P(x) \rightarrow x = a \vee x = d)$ then $D(\mathcal{V}(P(x)), C(P(x))) = \{< a >, < d >\}$

$\forall x(Q(x) \rightarrow x = b \vee x = c)$ then $D(\mathcal{V}(Q(x)), C(Q(x))) = \{< b >, < c >\}$

$$\forall x, y(R(x,y) \rightarrow (x = a \wedge y = b) \vee (x = a \wedge y = c) \vee (x = b \wedge y = e)) \\ \text{then} \quad D(\mathcal{V}(R(x,y)), C(R(x,y))) = \{< a, b >, < a, c >, < b, e >\} \ .$$

If we have:

$\varphi_1 = P(x) \vee Q(x)$ then $D(\mathcal{V}(\varphi_1), C(\varphi_1)) = \{< a >, < b >, < c >, < d >\}$

$\varphi_2 = R(x,y) \wedge P(x)$ then $D(\mathcal{V}(\varphi_2), C(\varphi_2)) = \{< a, b >, < a, c >\}$ .

$\varphi_3 = R(x,y) \wedge \neg P(x)$ then $D(\mathcal{V}(\varphi_3), C(\varphi_3)) = \{< b, e >\}$ .

# Translation Procedure

## Quantifier elimination procedure

- if $D(\mathcal{V}(\varphi_1), C(\varphi_1)) = \{< \overline{c_1} >, ..., < \overline{c_n} >\}$ with $n > 0$:

$$T(\forall\overline{x}(\varphi_1(\overline{x}) \rightarrow \varphi_2(\overline{x})), C(\varphi)) = T(\varphi_2(\overline{c_1}), C(\varphi_2(\overline{c_1}))) \wedge ... \wedge T(\varphi_2(\overline{c_n}), C(\varphi_2(\overline{c_n}))) \ .$$

$$T(\exists\overline{x}(\varphi_1(\overline{x}) \wedge \varphi_2(\overline{x})), C(\varphi)) = T(\varphi_2(\overline{c_1}), C(\varphi_2(\overline{c_1}))) \vee ... \vee T(\varphi_2(\overline{c_n}), C(\varphi_2(\overline{c_n}))) \ .$$

- if $D(\mathcal{V}(\varphi_1), C(\varphi_1)) = \varnothing$ :

$$T(\forall\overline{x}\ (\varphi_1(\overline{x}) \rightarrow \varphi_2(\overline{x}))\ ,\ C(\varphi)) = True \ .$$

$$T(\exists\overline{x}\ (\varphi_1(\overline{x}) \wedge \varphi_2(\overline{x}))\ ,\ C(\varphi)) = False \ .$$

# Translation Procedure - Observation

## Observation 1

Let $F$ be a restricted formula of the form $F : \exists x(\varphi(x) \wedge \psi(x))$ where $\varphi$ is a domain formula, and its corresponding completion formula
$C(\varphi) : \forall x(\varphi(x) \leftrightarrow x = c_1 \vee x = c_2 \vee ... \vee x = c_n)$ .

Then we have:
$F' : \exists x((x = c_1 \vee x = c_2 \vee ... \vee x = c_n) \wedge \psi(x))$ .

Using the equality substitution axiom scheme we can prove that $F \leftrightarrow F''$ where:
$F'' : \psi(c_1) \vee ... \vee \psi(c_n)$ .

# Translation Procedure - Observation

## Observation 2

Let $F$ be a restricted formula of the form $F : \forall x(\varphi(x) \rightarrow \psi(x))$ where $\varphi$ is a domain formula, and its corresponding completion formula
$C(\varphi) : \forall x(\varphi(x) \leftrightarrow x = c_1 \vee x = c_2 \vee ... \vee x = c_n)$ .

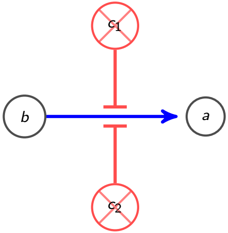
Then we have:
$F' : \forall x((x = c_1 \vee x = c_2 \vee ... \vee x = c_n) \rightarrow \psi(x))$ .

P
Using the equality substitution axiom scheme we can prove that $F \leftrightarrow F''$ where:
$F'' : \psi(c_1) \wedge ... \wedge \psi(x_n)$ .

# Example

$$\forall x(\exists y(A(y) \wedge CA(y,x) \wedge \forall z(CICA(z,y,x) \rightarrow \neg A(z))) \rightarrow A(x)) \quad (20)$$

- $\forall y(CA(y,a) \leftrightarrow y = b)$
- $\forall z(CICA(z,b,a) \leftrightarrow z = c_1 \vee z = c_2)$




$$A(b) \wedge \neg A(c_1) \wedge \neg A(c_2) \rightarrow A(a)$$

# Example - Continued

$$\forall x(\exists y(A(y) \wedge CA(y,x) \wedge \forall z(CICA(z,y,x) \rightarrow \neg A(z))) \rightarrow A(x))$$

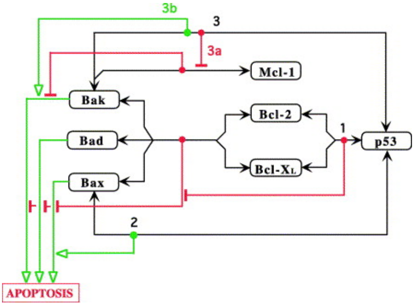
1. Our restricted formula is of the following form:
   $\exists y(CA(y,x) \wedge \varphi(y))$
   We can then apply the translation procedure using our first completion formula, thus eliminating $y$:

$$A(b) \wedge \forall z(CICA(z,b,a) \rightarrow \neg A(z)) \rightarrow A(a) \tag{21}$$

2. We can also apply a second translation procedure to $\forall z(CICA(z,b,a) \rightarrow \neg A(z))$ using the second completion formula, thus eliminating $z$. Which finally gives us:

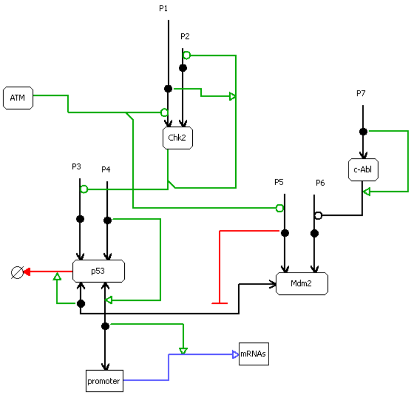
$$A(b) \wedge \neg A(c_1) \wedge \neg A(c_2) \rightarrow A(a) \tag{22}$$

# Example - Mitochondrial apoptosis induced by p53 independently of transcription




- $A(p53) \wedge A(bak) \rightarrow A(bak\_p53)$.
- $A(bak\_p53) \rightarrow I(bak\_mcl)$.
- $A(bak\_p53) \wedge \neg A(b\_complex) \wedge \neg A(bak\_mcl) \rightarrow A(apoptosis)$.
- $A(bak) \wedge \neg A(b\_complex) \wedge \neg A(bak\_mcl) \rightarrow A(apoptosis)$
- $A(p53) \wedge A(bcl) \rightarrow A(p53\_bb\_complex)$
- $A(p53\_bb\_complex) \rightarrow I(b\_complex)$
- $A(bax) \wedge \neg A(b\_complex) \rightarrow A(apoptosis)$
- $A(p53) \wedge A(bax) \wedge \neg A(b\_complex) \rightarrow A(apoptosis)$
- $A(bad) \wedge \neg A(b\_complex) \rightarrow A(apoptosis)$

# Example - DNA Double-Strand Break




- $A(atm) \wedge A(chk2) \rightarrow A(chk2\_p1)$
- $A(chk2) \wedge A(chk2\_p1) \rightarrow A(chk2\_p2)$
- $A(atm) \wedge A(mdm2) \rightarrow A(mdm2\_p5)$
- $A(c\_abl\_p7) \wedge A(mdm2) \rightarrow A(mdm2\_p6)$
- $A(chk2\_p2) \wedge A(p53) \rightarrow A(p53\_p3)$
- $A(p53) \wedge A(mdm2) \wedge \neg A(mdm2\_p5) \rightarrow A(mdm2\_p53)$
- $A(p53\_mdm2) \rightarrow A(p53\_degradation)$
- $A(p53\_p4) \wedge A(promoter) \rightarrow A(p53\_promoter)$
- $A(promoter) \wedge A(p53\_promoter) \rightarrow A(mrnas)$

# Example

## Questions

Abduction, Deduction

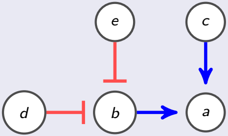
## Abduction for: $A(apoptosis)$

- $A(p53) \wedge A(bcl) \wedge A(bak)$: is a plausible answer, because p53 can bind to Bcl giving the *p53_bb_complex*, which can in return inhibit the *b_complex* that is responsible of inhibiting the capacity of Bak to activate the cell's apoptosis.
- Another interpretation of the previous answer is that p53 can also bind to Bak giving the *bak_p53* protein, which can in return inhibit the *bak_mcl* responsible of inhibiting the capacity of Bak to activate the cell's apoptosis. *bak_p53* can also stimulate Bak to reach apoptosis. Without forgetting that *p53_bb_complex* inhibit *b_complex*.
- $A(p53) \wedge A(bcl) \wedge A(bax)$: can also be a plausible answer.
- ...

# Another type of questions

## Test basis

- For $A(a)$ we should have $A(b)$ or $A(c)$.
- Consistency conditions for $A(b)$: $\neg A(e)$ and $\neg A(d)$.
- If we know that either $A(d)$ or $A(e)$, then we also know that only $c$ will activate $a$.



# Finally

## Possible extensions

- Quantities, concentrations...
- Time, order...
- Notion of Aboutness

Thank you.