Formal Molecular Biology

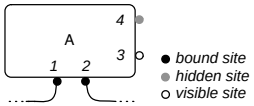
Presentation by Jérôme Caffaro & Jean-Philippe Pellet

Reference:
Vincent Danos & Cosimo Laneve,
Formal Molecular Biology,
http://www.cs.unibo.it/∼laneve/papers/fmb.pdf

1 A visual notation for κ-calculus

Definitions A *site* is a point of connection on a protein that can be linked to another protein.
Depending on how the protein is folded in space, a site can be active or not:

- A site may be *bound* or *free*.
- A free site can be *visible* or *hidden*.



Proteins can be assembled (connected) into *complexes*.

Definitions Collections of proteins and complexes are called *solutions*.
Solutions evolve by mean of *reactions*, which occurs when a sub-solution has a special shape, called a *reactant*.

2 κ-calculus

Definitions The syntax of κ-calculus relies on

- a countable set of *protein names* $\mathcal{P}$, ranged over by $A, B, C, \ldots$
- a countable set of *edge names* $\mathcal{E}$, ranged over by $x, y, z, \ldots$
- a *signature map*, written $\mathtt{s}$, from $\mathcal{P}$ to natural number $\mathbb{N}$.

For each protein, $\mathtt{s}(A)$ is the number of sites of A and the pair (A, i) is a *site* of A.

Definition (Interface) An *interface* is a partial map from $\mathbb{N}$ to $\mathcal{E} \cup \{\mathrm{h},\mathrm{v}\}$ ranged over by $\rho, \sigma, \ldots$
A site (A, i) is said to be:

- *visible* if $\rho(i) = v$
- *hidden* if $\rho(i) = h$
- *bound* if $\rho(i) \in \mathcal{E}$

Interfaces are used to depict partial states of A's sites.

Syntax The syntax is given by:

$\mathcal{S}$:=		**solution**
	0	empty solution
	$A(\rho)$	protein
	$\mathcal{S}, \mathcal{S}$	group
	$(\nu x)(\mathcal{S})$	new

- The "new" operator is a binder: in $(\nu x)(\mathcal{S})$, $\mathcal{S}$ is the *scope* of the binder (νx)
- We inductively define the set $\mathrm{fn}(\mathcal{S})$ of *free names* in a solution $\mathcal{S}$:

$$\begin{array}{rcl} \mathrm{fn}(0) & = & \emptyset \\ \mathrm{fn}(A(\rho)) & = & \mathrm{fn}(\rho) \\ \mathrm{fn}(\mathcal{S}, \mathcal{S}') & = & \mathrm{fn}(\mathcal{S}) \cup \mathrm{fn}(\mathcal{S}') \\ \mathrm{fn}((\nu x)(\mathcal{S})) & = & \mathrm{fn}(\mathcal{S}) \setminus \{x\} \end{array}$$

- An occurrence of x in $\mathcal{S}$ is *bound* if it occurs in a sub-solution which is in the scope of the binder x.
- A solution $\mathcal{S}$ is *closed* if all occurrences of names in $\mathcal{S}$ are bound ($\approx$ if $\mathrm{fn}(S) = \emptyset$).

Abbreviation: $(\nu x_1, ..., x_n)(\mathcal{S})$ or $(\nu \tilde{x})(\mathrm{S})$ instead of $(\nu x_1)...(\nu x_n)(\mathcal{S})$.

Definition (Structural Congruence) Structural congruence, written $\equiv$, is the least equivalence closed under syntactic conditions, containing α-equivalence (injective renaming of bound variables), taking "," to be associative (as the choice of symbols suggests) and commutative, with 0 as neutral element, and satisfying the scope laws:

$$\begin{array}{rcll} (\nu x)(\nu y)(\mathcal{S}) & \equiv & (\nu y)(\nu x)(\mathcal{S}) & \\ (\nu x)(\mathcal{S}) & \equiv & \mathcal{S} & \text{when } x \notin \mathrm{fn}(\mathcal{S}) \\ (\nu x)(\mathcal{S}), \mathcal{S}' & \equiv & (\nu x)(\mathcal{S}, \mathcal{S}') & \text{when } x \notin \mathrm{fn}(\mathcal{S}') \end{array}$$

Using structural congruence, we can define *connectedness*:

- $\mathcal{A}(\rho)$ is connected;
- if $\mathcal{S}$ is connected so is $(x)(\mathcal{S})$
- if $\mathcal{S}$ and $\mathcal{S}'$ are connected and $\mathrm{fn}(\mathcal{S}) \cap \mathrm{fn}(\mathcal{S}') \neq \emptyset$ then $\mathcal{S}, \mathcal{S}'$ is connected;
- if $\mathcal{S}$ is connected and $\mathcal{S} \equiv \mathcal{T}$ then $\mathcal{T}$ is connected.

Definition (Graph-likeness) A solution is said to be *graph-like* iff:

- free names occur at most twice in $\mathcal{S}$;
- binders in $\mathcal{S}$ bind either zero or two occurrences.

If in addition free names occurs exactly twice in $\mathcal{S}$, we say that $\mathcal{S}$ is *strongly graph-like*.

Definition (Graph with sites) Let $[\![.]\!]_g$ be the following function from graph-like solutions to graphs with sites:

- $[\![A(\rho)]\!]_g$ is the graph with a single node labeled A, sites in $\{1, \ldots, \mathfrak{s}(A)\}$, bound sites k being labeled by $\rho(k)$, and free sites being in the state prescribed by ρ
- $[\![\mathcal{S}, \mathcal{S}']\!]_g$ is the union graph of $[\![\mathcal{S}]\!]_g$ and $[\![\mathcal{S}']\!]_g$, with sites labeled with the same name being connected by an edge, and their common name erased
- $[\![(\nu x)(\mathcal{S})]\!]_g$ is $[\![\mathcal{S}]\!]_g$

Definition (Growth relation $\leq$) The *growth relation* $\leq$ is the least binary relation over interfaces precisely generated by the following rules:

$$\text{(CREATE)}: \ \frac{x \in \tilde{x}}{\tilde{x} \vdash \imath \leq \imath^x}$$

$$\text{(HV-SWITCH)}: \ \frac{}{\tilde{x} \vdash \bar{\imath} \leq \imath}$$

$$\text{(VH-SWITCH)}: \ \frac{}{\tilde{x} \vdash \imath \leq \bar{\imath}}$$

$$\text{(REFLEX)}: \ \frac{\tilde{x} \cap \mathrm{fn}(\rho) = \emptyset}{\tilde{x} \vdash \rho \leq \rho}$$

$$\text{(SUM)}: \ \frac{\tilde{x} \vdash \rho \leq \rho' \quad \tilde{x} \vdash \sigma \leq \sigma'}{\tilde{x} \vdash \rho + \sigma \leq \rho' + \sigma'}$$

Definition (Growth relation $\leq$ for pre-solutions) The *growth relation* $\leq$ is extended to the pre-solutions as the least binary relation precisely generated by the following rules:

$$\text{(NIL)}: \ \frac{}{\tilde{x} \vdash 0 \leq 0}$$

$$\text{(GROUP)}: \ \frac{\tilde{x} \vdash \mathcal{S} \leq \mathcal{S}' \quad \tilde{x} \vdash \rho \leq \rho' \quad \mathrm{dom}(\rho') \subseteq \mathfrak{s}(A)}{\tilde{x} \vdash \mathcal{S}, A(\rho) \ \leq \ \mathcal{S}', A(\rho')}$$

$$\text{(SYNTH)}: \ \frac{\tilde{x} \vdash \mathcal{S} \leq \mathcal{S}' \quad \mathrm{fn}(\rho) \subseteq \tilde{x} \quad \mathrm{dom}(\rho) = \mathfrak{s}(A)}{\tilde{x} \vdash \mathcal{S} \ \leq \ \mathcal{S}', A(\rho)}$$

Definition (Monotonic reaction) Let $\mathcal{S}, \mathcal{S}'$ be two pre-solutions. The reaction $\mathrm{r}_1 : \mathcal{S} \to (\nu\tilde{x})\mathcal{S}'$ is called *monotonic* iff:

- $\tilde{x} \vdash \mathcal{S} \leq \mathcal{S}'$
- $\mathcal{S}$ and $(\nu\tilde{x})\mathcal{S}'$ are graph-like
- $\mathcal{S}'$ is connected

Lemma $\mathrm{fn}(\mathcal{S}) = \mathrm{fn}((\nu\tilde{s})\mathcal{S}')$

Definition (Antimonotonic reaction) Let $\mathcal{S}, \mathcal{S}'$ be two pre-solutions. The reaction $\mathrm{r}_2 : (\nu\tilde{x})\mathcal{S} \to \mathcal{S}'$ is called *antimonotonic* iff its dual $\mathcal{S}' \to (\nu\tilde{x})\mathcal{S}$ is monotonic.

Lemma $\mathrm{fn}((\nu\tilde{x})\mathcal{S}) = \mathrm{fn}(\mathcal{S}')$

Definition (Renaming) A *renaming* r is a partial finite injection on $\mathcal{E} \cup \{h, v\}$, which is the identity on $\{h, v\}$ and maps $\mathcal{E}$ onto $\mathcal{E}$.

Definition (Matching solutions (monotonic)) Let $\mathcal{R} \to (\nu\tilde{x})\mathcal{P}$ be a monotonic reaction, and $\mathcal{S}, \mathcal{T}$ be two solutions.

$\mathcal{S}, \mathcal{T}$ match $\mathcal{R} \to (\nu\tilde{x})\mathcal{P}$

$\Leftrightarrow$ $\mathcal{S}, \mathcal{T} \models \mathcal{R} \to (\nu\tilde{x})\mathcal{P}$

$\Leftrightarrow$ $\exists$ renaming r, $\exists$ partial interfaces $\xi_1, \cdots, \xi_m$ such that:

1. $\forall i : r(\tilde{x}) \cap \mathrm{fn}(\xi_i) = \emptyset$
2. $\mathcal{S} = A_1(r \circ \rho_1 \cup \xi_1), \cdots, A_n(r \circ \rho_n \cup \xi_n)$

and $\mathcal{T} = \big(r(\tilde{x})\big)\big(A_1(r \circ \sigma_1 \cup \xi_1), \cdots, A_m(r \circ \sigma_m \cup \xi_m)\big)$

Definition (Matching solutions (antimonotonic)) Let $(\nu\tilde{x})\mathcal{R} \to \mathcal{P}$ be a monotonic reaction, and $\mathcal{S}, \mathcal{T}$ be two solutions.

$\mathcal{S}, \mathcal{T}$ match $(\nu\tilde{x})\mathcal{R} \to \mathcal{P}$

$\Leftrightarrow$ $\mathcal{S}, \mathcal{T} \models (\nu\tilde{x})\mathcal{R} \to \mathcal{P}$

$\Leftrightarrow$ $\mathcal{T}, \mathcal{S} \models \mathcal{P} \to (\nu\tilde{x})\mathcal{R}$

Definition (Transition relation $\to_{\mathrm{R}}$) Given a set of biological reactions R and a set of solutions, the *transition relation* $\to_{\mathrm{R}}$ is the least binary relation over solutions precisely generated by the following rules:

$$\text{(MON)}: \ \frac{\mathcal{S}, \mathcal{T} \models \mathcal{R} \to (\nu\tilde{x})\mathcal{P} \ \in \mathrm{R}}{\mathcal{S} \to_{\mathrm{R}} \mathcal{T}}$$

$$\text{(ANTIMON)}: \ \frac{\mathcal{S}, \mathcal{T} \models (\nu\tilde{x})\mathcal{R} \to \mathcal{P} \ \in \mathrm{R}}{\mathcal{S} \to_{\mathrm{R}} \mathcal{T}}$$

$$(\text{NEW}):\ \frac{\mathcal{S} \to_{\mathrm{R}} \mathcal{T}}{(\nu x)\mathcal{S} \to_{\mathrm{R}} (\nu x)\mathcal{T}}$$

$$(\text{GROUP}):\ \frac{\mathcal{S} \to_{\mathrm{R}} \mathcal{T}}{\mathcal{S},\mathcal{U} \to_{\mathrm{R}} \mathcal{T},\mathcal{U}}$$

$$(\text{STRUCT}):\ \frac{\mathcal{S} \to_{\mathrm{R}} \mathcal{T} \quad \mathcal{S} \equiv \mathcal{S}' \quad \mathcal{T} \equiv \mathcal{T}'}{\mathcal{S}' \to_{\mathrm{R}} \mathcal{T}'}$$

3 $m\kappa$-calculus

Definition (Extended interface) An *extended interface* θ is a map from $\mathbb{N}$ to $(\mathcal{E} \cup \mathcal{G} \cup \{h, v\}) \times \mathbb{N}$.

Definition (Projection of extended interfaces) An extended interface θ from $m\kappa$ can be projected to standard interface $[\theta]^-$ from κ with the following map:

- $[\imath^{x,n}]^- = \imath^x$
- $[\imath^{r,n}]^- = [\imath^{v,n}]^- = [\imath^n]^- = \imath^v = \imath$
- $[\imath^{h,n}]^- = [\bar{\imath}^n]^- = \imath^h = \bar{\imath}$

Projection is thus extended to interfaces, agents and solutions.

Definition ((Anti)monotonic interaction) With $\mathcal{R}, \mathcal{P}$ two pre-solutions, $\mathcal{R} \to \mathcal{P}$ is a *monotonic* (resp. an *antimonotonic*) *interaction* iff:

1. $\mathcal{R}$ and $\mathcal{P}$ consist of at most two agents
2. $\mathrm{fn}(\mathcal{R}) \supseteq \mathrm{fn}(\mathcal{P})$
3. $\mathrm{bn}(\mathcal{R}) \cap \mathcal{G} = \emptyset$
4. its projection $[\mathcal{R}]^- \to [\mathcal{P}]^-$ is monotonic (resp. antimonotonic) in κ

Definition (Micro-scenario) A *micro-scenario* for a monotonic reaction $\mathrm{r} : \mathcal{R} \to (\nu\tilde{x})\mathcal{P}$ is a tuple $(\mathcal{F}_{\mathrm{r}}, \mathcal{T}_{\mathrm{r}}, \mathrm{init})$, where:

- $\mathcal{F}_{\mathrm{r}}$ is a directed acyclic version of $[\![\mathcal{P}]\!]_g$ (the graph of the products)
- $\mathcal{T}_{\mathrm{r}}$is a tree spanning the flow graph $\mathcal{F}_{\mathrm{r}}$
- init is the common root of both $\mathcal{F}_{\mathrm{r}}$ and $\mathcal{T}_{\mathrm{r}}$

Definition (Signal Ordering Relation $\succ$) The *signal ordering relation* $\succ$ is the least transitive relation over extended sites such that:

$$\begin{aligned} \mathcal{F}_{\mathrm{r}}(a,i) = (b,j) \quad &\Rightarrow \quad (a,i) \succ (b,j) \\ \mathcal{F}^*_{\mathrm{r}}(a,i) \neq \bot \wedge \mathcal{F}_{\mathrm{r}}(a,j) \neq \bot \quad &\Rightarrow \quad (a,i) \succ (a,j) \end{aligned}$$

Notation (IN and OUT interfaces) With $\tilde{x} = (x_1, x_2, \cdots, x_k)$:

$$\mathrm{IN}_a^{\tilde{x},n} \stackrel{\text{def}}{=} \bigcup_{\{i \mid \mathcal{F}_r^*(a,i) \neq \bot\}} \imath^{x_i,n}$$

$$\mathrm{OUT}_a^{\tilde{x},n} \stackrel{\text{def}}{=} \bigcup_{\{i \mid \mathcal{F}_r(a,i) \neq \bot\}} \imath^{x_i,n}$$

Definition (The Monotonic Protocol) Given a monotonic reaction $\mathcal{R} \to (\nu\tilde{x})(\mathcal{P})$, the corresponding transitions in $m\kappa$ are precisely generated by the following rules:

Initiation and first contacts:

$$(\text{INIT}):\ \frac{a = \mathrm{init}(\mathcal{F}_{\mathrm{r}})}{A(\sigma) \to (\nu r)(A^{r,a}(\sigma'))}$$

$$(\text{FC}_1):\ \frac{\mathcal{T}_{\mathrm{r}}(a,i) = (b,j) \quad x \in \mathrm{fn(r)}}{A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i^x), B(j^x + \sigma) \to A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i^{x,1}), B^{r,b}(j^{x,1} + \sigma')}$$

$$(\text{FC}_2):\ \frac{\mathcal{T}_{\mathrm{r}}(a,i) = (b,j) \quad x \notin \mathrm{fn(r)} \quad b \in \mathcal{R}}{A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i), B(j + \sigma) \to (\nu x)\big(A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i^{x,1}), B^{r,b}(j^{x,1} + \sigma')\big)}$$

$$(\text{FC}_3):\ \frac{\mathcal{T}_{\mathrm{r}}(a,i) = (b,j) \quad x \notin \mathrm{fn(r)} \quad b \notin \mathcal{R}}{A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i) \to (\nu x)\big(A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i^{x,1}), B^{r,b}(j^{x,1} + \sigma)\big)}$$

Later contacts and responses:

$$(\text{LC}_1):\ \frac{\mathcal{T}^c_{\mathrm{r}}(a,i) = (b,j) \quad x \in \mathrm{fn(r)}}{A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i^x), B^{r,b}(j^x) \to A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i^{x,1}), B^{r,b}(j^{x,1})}$$

$$(\text{LC}_2):\ \frac{\mathcal{T}^c_{\mathrm{r}}(a,i) = (b,j) \quad x \notin \mathrm{fn(r)}}{A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i^x), B^{r,b}(j) \to (\nu x)\big(A^{r,a}(\mathrm{IN}_a^{\tilde{y},1} + i^{x,1}), B^{r,b}(j^{x,1})\big)}$$

$$(\text{R}):\ \frac{\mathcal{F}_{\mathrm{r}}(a,i) = (b,j)}{A^{r,a}(i^{x,1}), B^{r,b}(j^{x,1} + \mathrm{OUT}_b^{\tilde{y},2}) \to A^{r,a}(i^{x,2}), B^{r,b}(j^{x,2} + \mathrm{OUT}_b^{\tilde{y},2})}$$

Completions:

$$(\text{SHIFT}):\ \frac{a = \mathrm{init}(\mathcal{F}_{\mathrm{r}})}{A^{r,a}(\mathrm{OUT}_a^{\tilde{y},2}) \to A^{r,a}(\mathrm{OUT}_a^{\tilde{y},3})}$$

$$(\text{I-PPG}):\ \frac{a = \mathrm{init}(\mathcal{F}_{\mathrm{r}}) \quad \mathcal{F}_{\mathrm{r}}(a,i) = (b,j)}{A^{r,a}(i^{x,3}), B^{r,b}(j^{x,2}) \to A^{r,a}(i^{x,4}), B^{r,b}(j^{x,3})}$$

$$(\text{PPG}):\ \frac{a \neq \mathrm{init}(\mathcal{F}_{\mathrm{r}}) \quad \mathcal{F}_{\mathrm{r}}(a,i) = (b,j)}{A^{r,a}(\mathrm{IN}_a^{\tilde{y},3} + i^{x,2}), B^{r,b}(j^{x,2}) \to A^{r,a}(\mathrm{IN}_a^{\tilde{y},3} + i^{x,3}), B^{r,b}(j^{x,3})}$$

$$(\text{I-EXIT}):\ \frac{a = \mathrm{init}(\mathcal{F}_{\mathrm{r}})}{A^{r,a}(\mathrm{OUT}_a^{\tilde{x},4}) \to A(o_a^{\tilde{x}})}$$

$$(\text{EXIT}):\ \frac{a \neq \mathrm{init}(\mathcal{F}_{\mathrm{r}})}{A^{r,a}(\mathrm{IN}_a^{\tilde{y},3} + \mathrm{OUT}_a^{\tilde{x},3}) \to A(\imath_a^{\tilde{y}} + o_a^{\tilde{x}})}$$