Handling Fuzzy Unification and Generalization of First-Order Terms over Similar Signatures with Non-Aligned Argument Positions and *Partial* Argument Position Maps

A Preliminary Overview

Hassan Aït-Kaci   Gabriella Pasi

June 2018

- Please beware: **this presentation is a set of working notes that have not yet been thoroughly formally refined**

- **We extend** our recent results on fuzzy $\mathcal{FOT}$ **unification and generalization** when signatures may have **similar pairs not involving *all* the arguments of either functors**

OK... And why should we care?...

- This applies in ***Fuzzy IR*** when database **records have no guarantee that the fields of a pair of similar objects are aligned nor that *all* contribute to the similarity in either side**

Our previous work

(we assume known all notions and notation defined there)

Recently, we presented 3 lattice structures over $\mathcal{FOT}$s (1 crisp and 2 fuzzy), gave declarative axioms and rules and expressed the 6 corresponding dual lattice operations as constraints:

- **Conventional signature**
  - • Unification (*Herbrand–Martelli&Montanari's*)
  - ✔ Generalization (*declarative version of Reynolds–Plotkin's*)
- **Signature with aligned similarity**
  - • "Weak" fuzzy unification (*Sessa's*)
  - ✔ "Weak" fuzzy generalization (*dual to Sessa's*)
- **Signature with misaligned similarity**
  - ✔ Full fuzzy unification (*different/mixed arities*)
  - ✔ Full fuzzy generalization (*different/mixed arities*)

(✔ indicates original contribution)

Unifying similar functors w/ different arg. number/order

GENERIC WEAK TERM DECOMPOSITION

$$\frac{(\,E \cup \{f(s_1,\,\cdots,s_m) \doteq g(t_1,\,\cdots,t_n)\}\,)_\alpha}{(\,E \cup \{s_1 \doteq t_{p(1)},\,\cdots,s_m \doteq t_{p(m)}\}\,)_{\alpha\wedge\beta}}$$

[s.t. $f \sim^p_\beta g;\ 0 \leq m \leq n$]

FUZZY EQUATION REORIENTATION

$$\frac{(\,E \cup \{f(s_1,\,\cdots,s_m) \doteq g(t_1,\,\cdots,t_n)\}\,)_\alpha}{(\,E \cup \{g(t_1,\,\cdots,t_n) \doteq f(s_1,\,\cdots,s_m)\}\,)_\alpha}$$

[s.t. $0 \leq n < m$]

Generalizing similar functors w/ different arg. number/order

Functor/Arity Similarity Left

$$\frac{\begin{pmatrix}\sigma_1^0\\ \sigma_2^0\end{pmatrix}_{\alpha_0} \vdash \begin{pmatrix}s_1'\\ t_1'\end{pmatrix} u_1 \begin{pmatrix}\sigma_1^1\\ \sigma_2^1\end{pmatrix}_{\alpha_1} \quad \cdots \quad \begin{pmatrix}\sigma_1^{m-1}\\ \sigma_2^{m-1}\end{pmatrix}_{\alpha_{m-1}} \vdash \begin{pmatrix}s_m'\\ t_m'\end{pmatrix} u_m \begin{pmatrix}\sigma_1^m\\ \sigma_2^m\end{pmatrix}_{\alpha_m}}{\begin{pmatrix}\sigma_1^0\\ \sigma_2^0\end{pmatrix}_{\alpha} \vdash \begin{pmatrix}f(s_1,\ldots,s_m)\\ g(t_1,\ldots,t_n)\end{pmatrix} f(u_1,\ldots,u_m) \begin{pmatrix}\sigma_1^m\\ \sigma_2^m\end{pmatrix}_{\alpha_m}}$$

[s.t. $f \sim_\beta^p g$; $0 \leq m \leq n$; $\alpha_0 \stackrel{def}{=} \alpha \wedge \beta$]

where, for $i = 1, \ldots, m$:

$$\begin{pmatrix}s_i'\\ t_i'\end{pmatrix}_{\beta_i} \stackrel{def}{=} \begin{pmatrix}s_i\\ t_{p(i)}\end{pmatrix}\uparrow_{\alpha_{i-1}} \begin{pmatrix}\sigma_1^{i-1}\\ \sigma_2^{i-1}\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}\sigma_1^{i-1}\\ \sigma_2^{i-1}\end{pmatrix}_{\beta_i} \vdash \begin{pmatrix}s_i'\\ t_i'\end{pmatrix} u_i \begin{pmatrix}\sigma_1^i\\ \sigma_2^i\end{pmatrix}_{\alpha_i}$$

Generalizing similar functors w/ different arg. number/order (ctd.)

FUNCTOR/ARITY SIMILARITY RIGHT

$$\frac{\begin{pmatrix}\sigma_1^0\\ \sigma_2^0\end{pmatrix}_{\alpha_0} \vdash \begin{pmatrix}s'_1\\ t'_1\end{pmatrix} u_1 \begin{pmatrix}\sigma_1^1\\ \sigma_2^1\end{pmatrix}_{\alpha_1} \quad \cdots \quad \begin{pmatrix}\sigma_1^{n-1}\\ \sigma_2^{n-1}\end{pmatrix}_{\alpha_{n-1}} \vdash \begin{pmatrix}s'_n\\ t'_n\end{pmatrix} u_n \begin{pmatrix}\sigma_1^n\\ \sigma_2^n\end{pmatrix}_{\alpha_n}}{\begin{pmatrix}\sigma_1^0\\ \sigma_2^0\end{pmatrix}_{\alpha} \vdash \begin{pmatrix}f(s_1,\ldots,s_m)\\ g(t_1,\ldots,t_n)\end{pmatrix} g(u_1,\ldots,u_n) \begin{pmatrix}\sigma_1^n\\ \sigma_2^n\end{pmatrix}_{\alpha_n}}$$

[s.t. $g \sim^p_\beta f;\;\; 0 \leq n \leq m;\;\; \alpha_0 \stackrel{\text{def}}{=} \alpha \wedge \beta$]

where, for $i = 1, \ldots, n$:

$$\begin{pmatrix}s'_i\\ t'_i\end{pmatrix}_{\beta_i} \stackrel{\text{def}}{=} \begin{pmatrix}s_{p(i)}\\ t_i\end{pmatrix}\Big\uparrow_{\alpha_{i-1}} \begin{pmatrix}\sigma_1^{i-1}\\ \sigma_2^{i-1}\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}\sigma_1^{i-1}\\ \sigma_2^{i-1}\end{pmatrix}_{\beta_i} \vdash \begin{pmatrix}s'_i\\ t'_i\end{pmatrix} u_i \begin{pmatrix}\sigma_1^i\\ \sigma_2^i\end{pmatrix}_{\alpha_i}$$

What about similar functors w/ only **partial** non-aligned arities?

E.g., $\mathtt{foo} \in \Sigma_5$ and $\mathtt{bar} \in \Sigma_4$ s.t. $\mathtt{foo} \approxeq \mathtt{bar}$

but where this similarity may be homomorphically extended from these functors to terms they construct *only* when:

- $\mathtt{foo}$'s 3^{rd} argument is similar to $\mathtt{bar}$'s 4^{th} argument
- $\mathtt{foo}$'s 4^{th} argument is similar to $\mathtt{bar}$'s 2^{nd} argument

I.e., $\exists$ **two mutually inverse** but **partial bijective mappings** between the argument positions of functors $\mathtt{foo}$ and $\mathtt{bar}$; *e.g.*,

- $\mu_{\mathtt{foo},\mathtt{bar}} : \{3, 4\} \to \{1, 2, 3, 4\} = \{3 \mapsto 4, 4 \mapsto 2\}$
- $\mu_{\mathtt{bar},\mathtt{foo}} : \{2, 4\} \to \{1, 2, 3, 4, 5\} = \{2 \mapsto 4, 4 \mapsto 3\}$

Similar functors w/ **partial** non-aligned arities (ctd.)

$\mu_{foo,bar}$

$\mu_{bar,foo}$

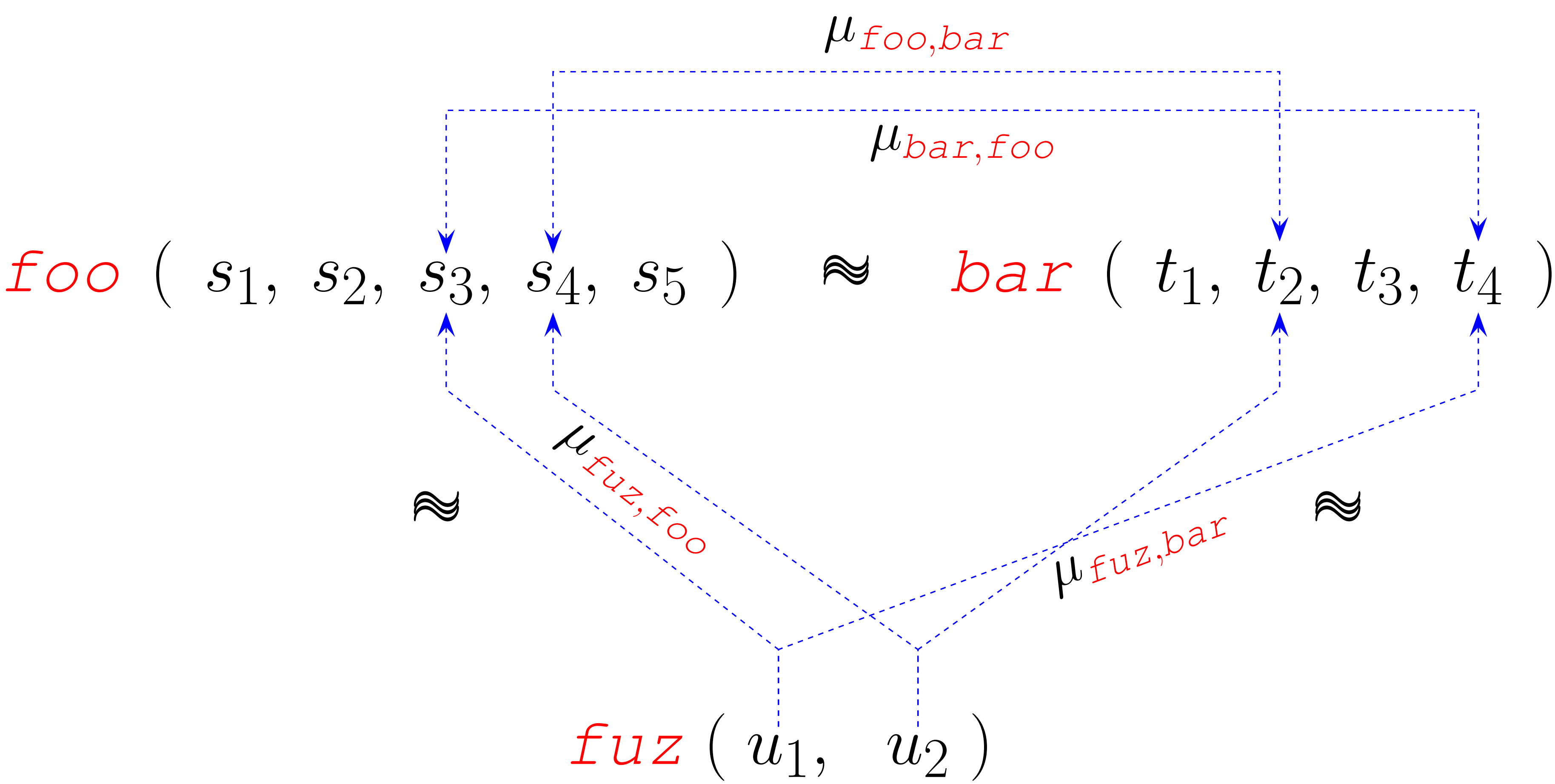
$$foo\,(s_1, s_2, s_3, s_4, s_5) \approx bar\,(t_1, t_2, t_3, t_4)$$

$\mu_{bar,foo}$

$\mu_{foo,bar}$

$foo \approx^{\mu_{foo,bar}} bar$ and $bar \approx^{\mu_{bar,foo}} foo$

Similar functors w/ **partial** non-aligned arities (ctd.)



$foo \approx^{\mu_{foo,bar}} bar$ and $bar \approx^{\mu_{bar,foo}} foo$

$fuz \approx^{\mu_{fuz,foo}} foo$ and $fuz \approx^{\mu_{fuz,bar}} bar$

Similar functors with **partial** non-aligned arities

Two functors $f \in \Sigma_m$ and $g \in \Sigma_n$, for any $m \geq 0$ and $n \geq 0$, are said to have *partial* non-aligned arities at approximation degree $\alpha \in (0, 1]$ whenever:

1. there is a set $\mathcal{D}^\alpha_{fg} \subseteq \{1, \ldots, m\}$ of argument positions of f and a set $\mathcal{D}^\alpha_{gf} \subseteq \{1, \ldots, n\}$ of argument positions of g such that $|\mathcal{D}^\alpha_{fg}| = |\mathcal{D}^\alpha_{gf}|$; and,
2. there exist two mutually inverse bijections:

$$\begin{cases} \mu^\alpha_{fg} : \mathcal{D}^\alpha_{fg} \to \{1, \ldots, n\} \\ \mu^\alpha_{gf} : \mathcal{D}^\alpha_{gf} \to \{1, \ldots, m\} \end{cases}$$

such that:

$$\begin{cases} \boldsymbol{ran}(\mu^\alpha_{fg}) = \mathcal{D}^\alpha_{gf} \stackrel{\text{def}}{=} \boldsymbol{dom}(\mu^\alpha_{gf}) \\ \boldsymbol{ran}(\mu^\alpha_{gf}) = \mathcal{D}^\alpha_{fg} \stackrel{\text{def}}{=} \boldsymbol{dom}(\mu^\alpha_{fg}) \end{cases}$$

Consistency conditions for **total** non-aligned arities (recall)

- for each $\langle f, g\rangle \in \Sigma^2$, s.t. $f \in \Sigma_m$ and $g \in \Sigma_n$, with $0 \leq m \leq n$, and $f \approx g$, there is an injective (*i.e.*, one-to-one) map $\mu_{fg} : \{1, \ldots, m\} \rightarrow \{1, \ldots, n\}$ associating each of the m argument positions of f to a unique position among the n arguments of g — this is denoted as: $f \approx^{\mu_{fg}} g$
- alignment maps between similar functors must be consistent:
  - for any functor f/n:

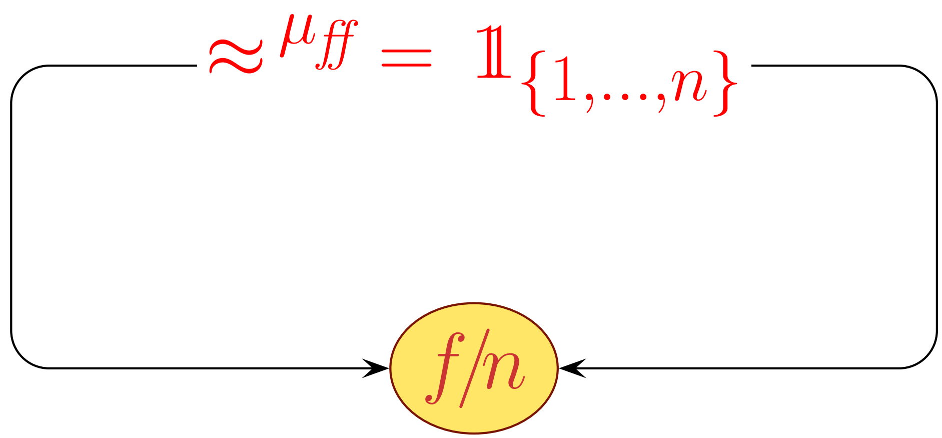
    Identity Consistency: $\mu_{ff} = \mathbb{1}_{\{1,\ldots,n\}}$
  - for any two functors f/n and g/n:

    Inverse Consistency: $\mu_{fg} \circ \mu_{gf} = \mathbb{1}_{\{1,\ldots,n\}}$
  - for any three functors f/m, g/n, h/ℓ s.t. $0 \leq m \leq n \leq \ell$:

    Composition Consistency: $\mu_{fh} = \mu_{gh} \circ \mu_{fg}$

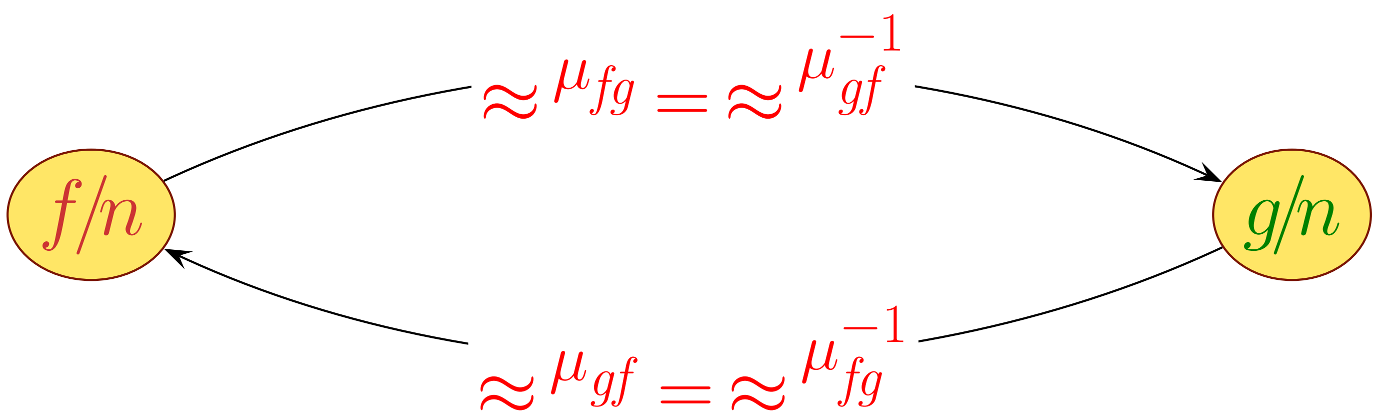
Consistency conditions for **total** non-aligned arities (recall-ctd.)

Identity Consistency Condition



Inverse Consistency Condition

$$\mu_{fg} : \{1, \ldots, n\} \to \{1, \ldots, n\}$$



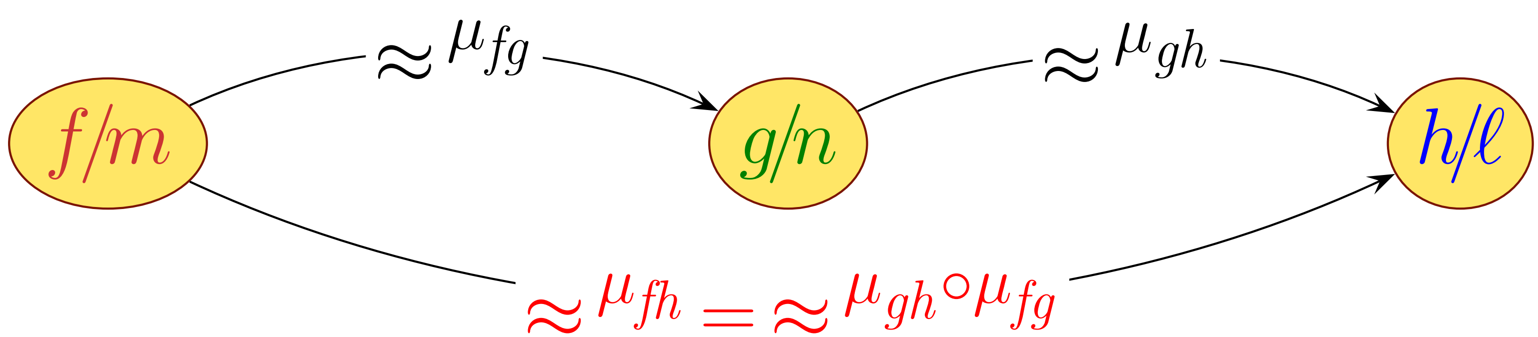
$$\mu_{gf} : \{1, \ldots, n\} \to \{1, \ldots, n\}$$

Consistency conditions for **total** non-aligned arities (recall-ctd.)

Compositional Consistency Condition

$$m \leq n \qquad\qquad n \leq \ell$$

$$\mu_{fg} : \{1, \ldots, m\} \to \{1, \ldots, n\} \qquad\qquad \mu_{gh} : \{1, \ldots, n\} \to \{1, \ldots, \ell\}$$



$$m \leq \ell$$

$$\mu_{fh} = \mu_{gh} \circ \mu_{fg} : \{1, \ldots, m\} \to \{1, \ldots, \ell\}$$

Consistency issues for **partial** non-aligned arities

***Total* argument-position maps are *always* composable** as *all* the positions in the range of a map are in the domain of any map from a functor to a similar one (always of greater or equal arity).

With *partial* maps, this may no longer be possible!

$\mathtt{foo}/5$, $\mathtt{bar}/4$, $\mathtt{biz}/4$, with:

$\mu^{\alpha}_{\mathtt{foo},\mathtt{bar}} : \{3,4\} \rightarrow \{2,4\}$ and $\mu^{\alpha}_{\mathtt{bar},\mathtt{biz}} : \{1,3\} \rightarrow \{1,2\}$

not composable: $\boldsymbol{ran}(\mu^{\alpha}_{\mathtt{foo},\mathtt{bar}}) \cap \boldsymbol{dom}(\mu^{\alpha}_{\mathtt{bar},\mathtt{biz}}) = \emptyset$

Even if $\mu^{\alpha}_{\mathtt{bar},\mathtt{biz}} : \{2,4\} \rightarrow \{1,2\}$ but $\mu^{\alpha}_{\mathtt{foo},\mathtt{biz}} : \{3,4\} \rightarrow \{3,4\}$:

$\mu^{\alpha}_{\mathtt{bar},\mathtt{biz}} \circ \mu^{\alpha}_{\mathtt{foo},\mathtt{bar}} \neq \mu^{\alpha}_{\mathtt{foo},\mathtt{biz}}$ not consistent

These are clearly situations to be detected: **partial argument maps must always be consistently composable**

Consistency conditions for **partial** non-aligned arities

Now, assume all $f \approxeq_\alpha g$ with partial non-aligned arities at approximation degree $\alpha \in (0,1]$ *(argument maps are identities by default)*

- for any $f \in \Sigma$, $g \in \Sigma$, $\alpha \in (0,1]$, $\beta \in (0,1]$:

$$\alpha \leq \beta \quad \Rightarrow \quad \mathcal{D}^{\alpha}_{fg} \subseteq \mathcal{D}^{\beta}_{fg}$$

- for any $f \in \Sigma$, $g \in \Sigma$, $\alpha \in (0,1]$, $\beta \in (0,1]$:

$$\alpha \leq \beta \quad \Rightarrow \quad \mu^{\alpha}_{fg} \subseteq \mu^{\beta}_{fg}$$

(as sets of pairs)

Consistency conditions for **partial** non-aligned arities (ctd.)

- for all $f \in \Sigma_m$, $g \in \Sigma_n$, $h \in \Sigma_\ell$, $m \geq 0$, $n \geq 0$, $\ell \geq 0$:

$$\begin{cases} \boldsymbol{ran}(\mu^\alpha_{fg}) = \boldsymbol{dom}(\mu^\alpha_{gh}) \;\; (= \mathcal{D}^\alpha_{gh}) \\ \boldsymbol{ran}(\mu^\alpha_{fh}) = \boldsymbol{ran}(\mu^\alpha_{gh}) \end{cases}$$

and:

composition order

$$\begin{cases} \mu^\alpha_{hf} = \mu^\alpha_{gf} \circ \mu^\alpha_{hg} \\ \mu^\alpha_{hg} = \mu^\alpha_{fg} \circ \mu^\alpha_{hf} \end{cases}$$

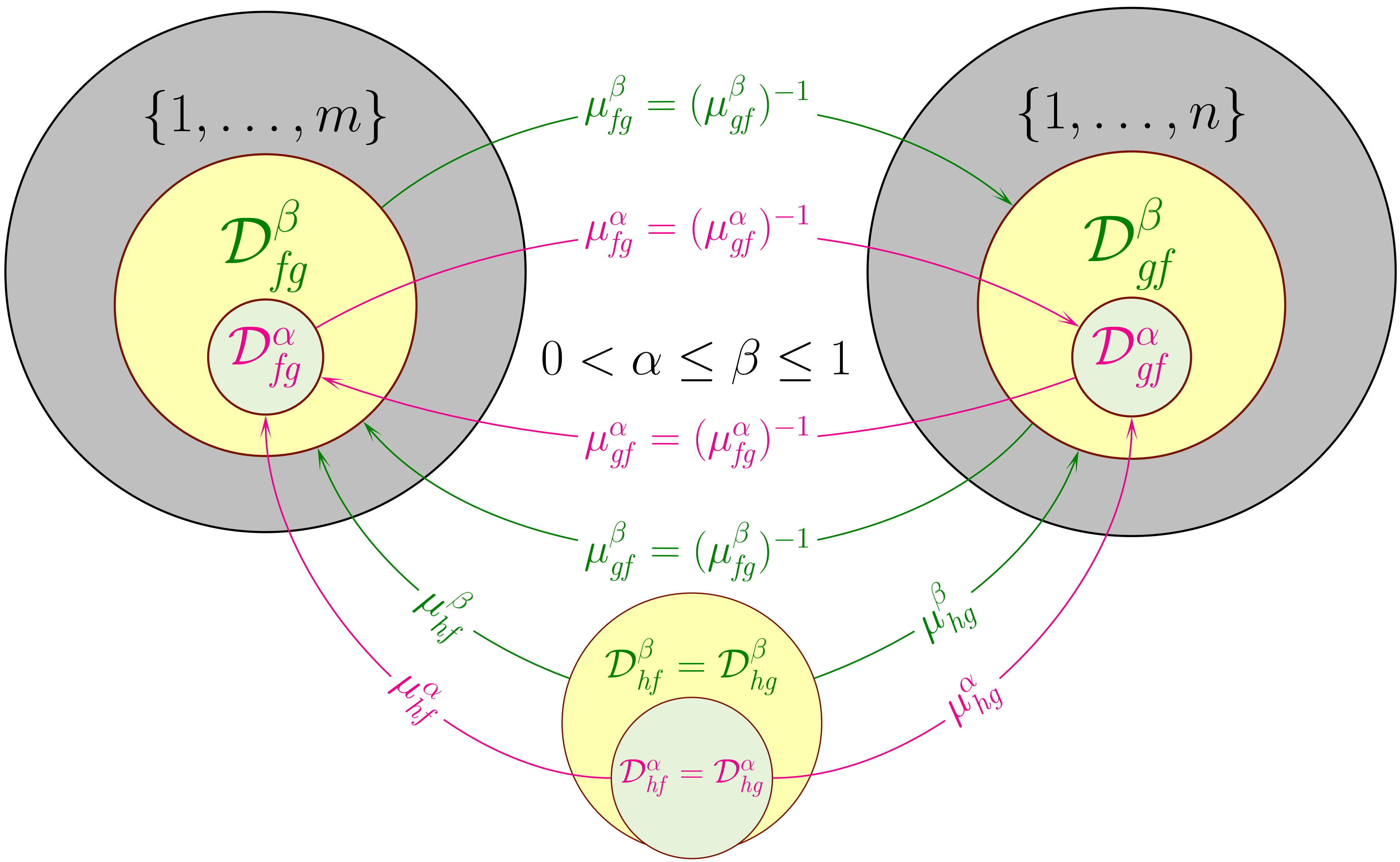
or

application order

$$\begin{cases} \mu^\alpha_{hf} = \mu^\alpha_{hg}\mu^\alpha_{gf} \\ \mu^\alpha_{hg} = \mu^\alpha_{hf}\mu^\alpha_{fg} \end{cases}$$

at any approximation degree $\alpha \in (0, 1]$

Partial non-aligned arity consistency as a commutative diagram



Unification with **partial** non-aligned arguments

PARTIAL NON-ALIGNED TERM DECOMPOSITION

$$\frac{(\, E \cup \{\, f(s_1, \ldots, s_m) \doteq g(t_1, \ldots, t_n) \,\}\,)_\alpha}{(\, E \cup \{\, s_{d_1} \doteq t_{\mu_{fg}^{\alpha\wedge\beta}(d_1)}, \ldots, s_{d_k} \doteq t_{\mu_{fg}^{\alpha\wedge\beta}(d_m)} \,\}\,)_{\alpha\wedge\beta}}$$

[s.t. $f \approx_{\beta}^{\mu_{fg}^{\beta}} g;\;\; 0 \leq |\mathcal{D}_{fg}^{\alpha\wedge\beta}| = k \leq \min(m,n);\;\; \mathcal{D}_{fg}^{\alpha\wedge\beta} = \{d_1, \ldots d_k\}$]

N.B.: there is no need to re-orient a term equation as for total maps! (**Why?**)

Generalization with **partial** non-aligned arguments

PARTIAL FUNCTOR/ARITY SIMILARITY

$$\frac{\begin{pmatrix}\sigma_1\\ \sigma_2\end{pmatrix}_{\alpha_0} \vdash \begin{pmatrix}s'_1\\ t'_1\end{pmatrix} u_1 \begin{pmatrix}\sigma_1^1\\ \sigma_2^1\end{pmatrix}_{\alpha_1} \quad \cdots \quad \begin{pmatrix}\sigma_1^{\ell-1}\\ \sigma_2^{\ell-1}\end{pmatrix}_{\alpha_{\ell-1}} \vdash \begin{pmatrix}s'_\ell\\ t'_\ell\end{pmatrix} u_\ell \begin{pmatrix}\sigma_1^\ell\\ \sigma_2^\ell\end{pmatrix}_{\alpha_\ell}}{\begin{pmatrix}\sigma_1\\ \sigma_2\end{pmatrix}_{\alpha} \vdash \begin{pmatrix}f(s_1,\ldots,s_m)\\ g(t_1,\ldots,t_n)\end{pmatrix} h(u_1,\ldots,u_\ell) \begin{pmatrix}\sigma_1^\ell\\ \sigma_2^\ell\end{pmatrix}_{\alpha_\ell}}$$

[s.t. $f/m \approx_\beta g/n;\ \ \alpha_0 \stackrel{def}{=} \alpha \wedge \beta;\ \ h/\ell \in [f/m, g/n]_{\alpha_0};\ \ |\mathcal{D}_{hf}^{\alpha_0}| = |\mathcal{D}_{hg}^{\alpha_0}| = \ell$]

where, for $i = 1,\ldots,\ell$:

$$\begin{pmatrix}s'_i\\ t'_i\end{pmatrix}_{\beta_i} \stackrel{def}{=} \begin{pmatrix}s_{\mu_{hf}^{\alpha_{i-1}}(i)}\\ t_{\mu_{hg}^{\alpha_{i-1}}(i)}\end{pmatrix} \uparrow_{\alpha_{i-1}} \begin{pmatrix}\sigma_1^{i-1}\\ \sigma_2^{i-1}\end{pmatrix} \text{ and } \begin{pmatrix}\sigma_1^{i-1}\\ \sigma_2^{i-1}\end{pmatrix}_{\beta_i} \vdash \begin{pmatrix}s'_i\\ t'_i\end{pmatrix} u_i \begin{pmatrix}\sigma_1^i\\ \sigma_2^i\end{pmatrix}_{\alpha_i}$$

N.B.: there is no need to differentiate between left and right as for total maps! (**Why?**)

But these rules work only if all consistency conditions hold!

Isn't similarity consistency a lot to ask?...

Most certainly!. . . However, **the good news is**:

- **an inconsistent signature similarity can easily be detected**
- **a consistent but incomplete signature can be completed** and remain consistent s.t. similarity classes always contain a least-arity functor with total consistent composable argument maps to all other members of its class

and **this can be done efficiently**!

Automated completion of **partial** non-aligned signature similarity

forall $\alpha \in \textsc{Val}^{\approx}$ and similarity class $c \in \Pi_{\alpha}^{\approx}$ **do**

if $\nexists$ a least-arity similarity class representative in c with *total* argument-position maps to *all* members of c

then

- add a *new* functor h/m to signature Σ_m such that $m = \min\{|\mathcal{D}_{fg}^{\alpha}| \mid f \in c,\ g \in c\}$, with least *total* consistent injective maps $\mu_{hf}^{\alpha} : \{1, \ldots, m\} \rightarrow \{1, \ldots, n\}$ for *all* $f/n \in c$; if not possible, Σ is inconsistent;
- add functor h/m to similarity class c;

Yet to be done. . .

☞ **Finish formal work**

– prove formal properties and correctness

– complete report and submit for publication

☞ **Java implementation**

– about 50% is done *(had to be put on hold to prepare this talk!)*

– yet to be done:

  ∗ partial map consistency checking

  ∗ automated least partial-map completion

☞ **Develop convincing examples!**

– use implementation to experiment on examples

☞ **HAK: find a job where they like this?** ☺, *etc.*, . . .

Thank You For Your Attention !

Hassan Aït-Kaci

hak@acm.org

Gabriella Pasi

pasi@disco.unimib.it