日语翻译的代数方法

VALENTIN BOBOC

Abstract. 我们采用 Lambek预群和语言的范畴组合分布模型（“DisCoCat”）框架，将日语到英语的翻译作为函子对进行研究。通过为预群添加装饰，我们展示了如何处理语言之间的语序变化。

arXiv:2303.05834v1 [cs.CL] 10 Mar 2023

1. 引言 Introduction

传统上，人们认为语言兼具经验性方面——人们通过练习语言来学习语言——以及组合性方面——复杂短语的意义完全由其结构及其组成部分的意义决定。

为了有效利用自然语言的组合性，一种广受欢迎的自然语言建模方法是范畴组合分布模型，简称“DisCoCat”（[CSC10]）。语言被建模为从解释语法（“组合性”）的范畴到解释语义（“分布式”）的范畴的函子。

组合部分负责判断短语或句子是否合乎语法：它根据各个组成部分的语法类型，计算出整个短语的语法类型。建模自然语言语法有多种代数方法。本文选择成熟的预群语法模型。预群由[Lam97]提出，旨在取代剩余幺半群的代数，以便对语法类型、它们的并置及约简进行建模。预群演算已被用于形式化表示多种自然语言的句法，例如：法语（[BL01a]）、德语（[LP04]）、波斯语（[Sad07]）、阿拉伯语（[BL01b]）、日语（[Car02]）和拉丁语（[CL05]）。

分布部分通过为单个词关联例如统计共现向量（[ML08]），为其赋予意义。因此，“DisCoCat”模型是一种通过语法结构解释意义组合的方式。

本文通过分析从日语到英语的翻译，探讨语言组合分布模型之间翻译的概念。在组合层面，翻译是一个强幺半函子。很容易证明，对于语序不同的语言，即便是翻译简单短语，这种函子也过于僵化，难以胜任。我们将说明，只要为底层预群语法

添加额外结构，就可以继续使用幺半函子这一工具。

我们首先介绍所采用的紧闭范畴的基本概念，即预群和有限生成向量空间，并定义翻译函子的概念。接下来，我们介绍日语语法的基础知识，以及用于为其建模的预群结构。最后，我们介绍预群装饰的概念，并以此构建一个用于翻译日语句子的结构化框架。

2. Theoretical background

2.1. 紧闭结构。“DisCoCat”模型的关键在于，预群范畴和有限生成向量空间范畴都是紧闭范畴。这使得语法的组合特性能够融入意义的分布空间中。

为完整起见，我们在此给出紧闭性的定义。建议读者参阅（[KL80]），以获取更全面且技术性更强的参考资料。

定义 1. 紧闭范畴是一个范畴 $\mathcal{C}$，并配有一个双函子

$$- \otimes - : \mathcal{C} \times \mathcal{C} \to \mathcal{C},$$

称为张量积，它在自然同构意义下满足结合律，并具有双侧单位元 I，且每个对象 $A \in \mathcal{C}$ 都有一个右对偶 A^r 和一个左对偶 A^ℓ，并带有以下态射

$$A \otimes A^r \xrightarrow{\varepsilon^r_A} I \xrightarrow{\eta^r_A} A^r \otimes A,$$

$$A^\ell \otimes A \xrightarrow{\varepsilon^\ell_A} I \xrightarrow{\eta^\ell_A} A \otimes A^\ell.$$

此外，ε 和 η映射满足“拉链”条件：

$$(1_A \otimes \varepsilon^\ell_A) \circ (\eta^\ell_A \otimes 1_A) = 1_A \qquad (\varepsilon^r_A \otimes 1_A) \circ (1_A \otimes \eta^r_A) = 1_A$$
$$(\varepsilon^\ell_A \otimes 1_A) \circ (1_{A^\ell} \otimes \eta^\ell_A) = 1_{A^\ell} \qquad (1_{A^r} \otimes \varepsilon^r_A) \circ (\eta^r_A \otimes 1_{A^r}) = 1_{A^r}.$$

紧闭性的要点是，我们希望有能够“相互抵消”的元素，并且可以将单位元分解为一个乘积。

2.2. 回顾预群。

定义 2. 预群是一个元组 $(P, \cdot, 1, -^\ell, -^r, \leqslant)$，其中 $(P, \cdot, 1, \leqslant)$ 是一个偏序幺半群，且一元运算 $-^\ell$、$-^r$（左对偶和右对偶）对所有 $x \in P$ 都满足以下关系：

$$x \cdot x^r \leqslant 1 \qquad x^\ell \cdot x \leqslant 1$$
$$1 \leqslant x^r \cdot x \qquad 1 \leqslant x \cdot x^\ell。$$

除非与上下文相关，否则省略运算符号 $\cdot$。不难验证，以下关系在每个预群中都成立：

$$1^\ell = 1 = 1^r \qquad (x^\ell)^r = x = (x^r)^\ell$$
$$(xy)^\ell = y^\ell x^\ell \text{ and } (xy)^r = y^r x^r \qquad \text{if } x \leqslant y \text{ then } y^\ell \leqslant x^\ell \text{ and } y^r \leqslant x^r.$$

我们通过从一组语法类型自由生成一个预群，来建模自然语言的语法。词典中的每个词都被赋予预群的一个元素，该元素对应于它的语言功能，例如名词、动词、形容词等。词串的解释方式是按照句法顺序，将赋予各组成部分的元素相乘。如果词串满足关系$w_1 w_2 \dots w_n \leqslant s$，我们就说该词串约化为类型 s。

例1。假设有两种语法类型：名词 n 和句子 s。语法建模为自由预群 $\mathbf{PGrp}(\{n, s\})$。考虑句子“鸽子吃面包。”我们将类型 n赋予“鸽子”和“面包”，并将类型$n^r s n^\ell$ 赋予及物动词“吃”。整个句子的类型为 $n(n^r s n^\ell)n$，并且有如下约化：

$$n(n^r s n^\ell)n = (nn^r)s(n^\ell n) \leqslant (1)s(n^\ell n) \leqslant s(1) \leqslant s.$$

在这种情况下，我们称“Pigeons eat bread”为合乎语法的句子，因为在预群 $\mathbf{PGrp}(\{n, s\})$ 中，这个短语可以化简为正确的类型。

这两个单独的化简可以按不同的顺序进行。Lambek 的 切换引理 [Lam97，命题 2] 告诉我们，在自由生成的预群中进行任何计算时，我们都可以不失一般性地假设所有收缩都先于所有扩张。

预群可以看作一个紧闭范畴。该范畴的对象是预群的元素。当且仅当$x \leqslant y$时，存在一个箭头 $x \to y$，而张量积由预群运算给出：$x \otimes y = xy$。态射 ε^r、ε^ℓ、η^r、η^ℓ 以显而易见的方式定义。用 ε 和 η 映射表示，本例中的化简可以表示为：

$$(\varepsilon^\ell_n \otimes 1_s \otimes \varepsilon^\ell_n)(n \otimes (n^r \otimes s \otimes n^\ell) \otimes n) \to s.$$

2.3. 意义空间。我们将自然语言的语义结构编码到有限生成向量空间范畴中，记作 $\mathbf{FVect}$。其箭头是线性变换，并且具有由线性代数张量积给出的自然幺结构，其单位元为 $\mathbb{R}$，而且该结构还是对称的：$V \otimes W \simeq W \otimes V$。这意味着 $V^\ell \simeq V^r \simeq V^*$，其中后者表示对偶向量空间。

为向量空间 V 固定一组基 $\{v_i\}$ 后，我们还可得到 $V \simeq V^*$，而紧闭性的结构态射由下式给出

$$\varepsilon_V = \varepsilon_V^r = \varepsilon_V^\ell : V^* \otimes V \to \mathbb{R} \text{ where } \sum_{i,j} a_{ij} v_i \otimes v_j \mapsto \sum_{i,j} a_{ij} \langle v_i | v_j \rangle$$

$$\eta_V = \eta_V^r = \eta_V^\ell : \mathbb{R} \to V \otimes V^* \text{ where } 1 \mapsto \sum_i v_i \otimes v_i \text{ extended linearly.}$$

如果我们用 P boththe 表示预群及其对应的范畴，那么语法与语义之间的桥由一个强幺半函子给出

$$F : P \to \mathrm{FVect},$$

我们称之为函子语言模型。该函子将向量空间赋给原子类型： $F(1) = I$、 $F(n) = N$ （名词的向量空间）、 $F(s) = S$ （句子的向量空间）等。对于 P 中的词语，幺半群性告诉我们 $F(x \otimes y) = F(x) \otimes F(y)$。紧闭性也得到保留： $F(x^\ell) = F(x^r) = F(x)^*$。例如，我们可以将类型为 $n^r s n^\ell$ 的及物动词“吃”解释为一个向量，位于

$$F(n^r \otimes s \otimes n^\ell) = F(n^r) \otimes F(s) \otimes F(n^\ell) = F(n)^* \otimes F(s) \otimes F(n)^* = N \otimes S \otimes N.$$

P 中的预群归约可以借助相应的 ε 和 η 映射解释为 FVect 中的语义归约。及物动词相关的归约因此由下式给出

$$F(\varepsilon_n^r \otimes 1_s \otimes \varepsilon_n^\ell) = F(\varepsilon_n^r) \otimes F(1_s) \otimes F(\varepsilon_n^\ell) = F(\varepsilon_n)^* \otimes F(1_s) \otimes F(\varepsilon_n)^* = \varepsilon_N \otimes 1_S \otimes \varepsilon_N.$$

句子或短语的意义，是通过将预群约化解释为短语中各个词的分布式意义的张量积所对应的语义约化而得出的。前面的示例“鸽子吃面包”可解释为

$$F(\varepsilon_n^r \otimes 1_s \otimes \varepsilon_n^\ell)(\mathbf{Pigeons} \otimes \mathbf{eat} \otimes \mathbf{bread}).$$

2.4. 函子语言模型之间的翻译。 （[BLMT18]）的作者对函子语言模型之间的翻译这一概念进行了形式化。我们通过一个简单名词短语翻译的示例，说明这一构造及可能遇到的问题。

定义 3. 设 $(\mathcal{C}, \otimes, 1_\mathcal{C})$ 和 $(\mathcal{D}, \odot, 1_\mathcal{D})$ 为幺半范畴。幺半函子 $F : \mathcal{C} \to \mathcal{D}$ 是一个函子，并配有自然同构 $\Phi_{x,\,y} : F(x) \odot F(y) \to F(x \otimes y)$，对每一对对象 x、 $y \in \mathcal{C}$ 均成立；还配有一个同构 $\phi : 1_\mathcal{D} \to F(1_\mathcal{C})$，使得对任意三个对象 x、 y、 $z \in \mathcal{C}$，以下图表可交换

$$\begin{array}{ccccc} (F(x) \odot F(y)) \odot F(z) & \xrightarrow{\Phi_{x,y} \odot 1_{F(z)}} & F(x \otimes y) \odot F(z) & \xrightarrow{\Phi_{x \otimes y, z}} & F((x \otimes y) \otimes z) \\ \downarrow & & & & \downarrow \\ F(x) \odot (F(y) \odot F(z)) & \xrightarrow{1_{F(x)} \odot \Phi_{y,z}} & F(x) \odot F(y \otimes z) & \xrightarrow{\Phi_{x, y \otimes z}} & F(x \otimes (y \otimes z)) \end{array}$$

其中，竖直箭头分别应用了各自范畴中的结合律。此外，对于每个对象 $x \in \mathcal{C}$，以下两个方块图都是交换的：

$$\begin{array}{ccc} 1_{\mathcal{D}} \odot F(x) & \longrightarrow & F(x) \\ \downarrow & & \uparrow \\ F(1_{\mathcal{C}}) \odot F(x) & \longrightarrow & F(1_{\mathcal{C}} \otimes x) \end{array} \qquad \begin{array}{ccc} F(x) \odot 1_{\mathcal{D}} & \longrightarrow & F(x) \\ \downarrow & & \uparrow \\ F(x) \odot F(1_{\mathcal{C}}) & \longrightarrow & F(x \otimes 1_{\mathcal{C}}). \end{array}$$

定义 4. 设 (F, Φ, ϕ) 和 (G, Ψ, ψ) 是幺半群范畴 $\mathcal{C}$ 与 $\mathcal{D}$ 之间的幺半群函子。幺半自然变换 $\alpha : F \Rightarrow G$ 是一种自然变换，使得以下图表交换：

$$\begin{array}{ccc} F(x) \odot F(y) & \xrightarrow{\alpha(x) \odot \alpha(y)} & G(x) \odot G(y) \\ \downarrow{\scriptstyle \Phi_{x,y}} & & \downarrow{\scriptstyle \Psi_{x,y}} \\ F(x \otimes y) & \xrightarrow{\alpha_{x \otimes y}} & G(x \otimes y) \end{array} \qquad \begin{array}{ccc} 1_{\mathcal{D}} & & \\ \downarrow{\scriptstyle \phi} & \searrow^{\psi} & \\ F(1_{\mathcal{C}}) & \xrightarrow{\alpha(1_{\mathcal{C}})} & G(1_{\mathcal{C}}). \end{array}$$

定义 5. 设 $\mathcal{A} : P \to \mathbf{FVect}$ 和 $\mathcal{B} : Q \to \mathbf{FVect}$ 是两个函子语言模型。从 F 到 G 的翻译是一个元组 (T, α)，其中 $T : P \to Q$ 是一个幺半函子，$\alpha : \mathcal{A} \Rightarrow \mathcal{B} \circ T$ 是一个幺半自然变换。

示例 2. 我们尝试将日语中形容词 $+$ 名词类型的简单短语翻译成英语。我们采用一个受限模型。设 $J = \mathbf{PGrp}(\{s_J, n_J\})$ 是由日语中的句子类型和名词类型生成的自由预群（或范畴），设 $E = \mathbf{PGrp}(\{s_E, n_E\})$ 是由英语中的句子类型和名词类型生成的自由预群。

函子语言模型分别记作 $\mathcal{J} : J \to \mathbf{FVect}$ 和 $\mathcal{E} : E \to \mathbf{FVect}$。语义赋值很直接：$\mathcal{J}(n_J) = N_J, \mathcal{J}(a_J) = A_J, \mathcal{E}(n_E) = N_E, \mathcal{E}(a_E) = A_E$。

该翻译将由以下幺半函子构成

$$T : J \to E,$$

它将 $s_J \mapsto s_E$ 和 $n_J \mapsto n_E$ 映射到相应对象。由此可知，相应语言中的类型约简得以保留，即 $T\left((n_J n_J^{\ell}) n_J\right) = T(n_J) = n_E$。由于幺半群性，只需定义自然变换 $\alpha : \mathcal{J} \Rightarrow \mathcal{E} \circ T$ 的分量 $\alpha_{n_J}, \alpha_{s_J}$，即可获得解析语义。

此外，自然变换 α 必须与幺半函子 T 交换。图示如下，这是一个交换方块：

$$\begin{array}{ccc} (N_J \otimes N_J) \otimes N_J & \xrightarrow{\mathcal{J}(\varepsilon^{\ell}_{N_J} \otimes 1_{\mathcal{J}})} & N_J \\ \downarrow{\scriptstyle \alpha_{(n_J n_J^{\ell}) n_J}} & & \downarrow{\scriptstyle \alpha_{n_J}} \\ (N_E \otimes N_E) \otimes N_E & \xrightarrow{\mathcal{E}(\varepsilon^{\ell} \otimes 1_{\mathcal{E}})} & N_E. \end{array}$$

考虑具体词语红色 $\in N_E \otimes N_E$、猫 $\in N_E$、akai $\in N_J \otimes N_J$ 和 neko $\in N_J$。该图表示：先运用日语语法规则将 akai⊗neko 化简为 akai neko，再翻译成 red cat，与先逐个翻译 akai ⊗neko，得到 red ⊗cat，再运用英语语法规则将其化简为 red cat，结果相同。

由于语序没有差异，这个短语翻译示例以预期的方式奏效。如果我们想把短语“akai neko”从日语翻译成罗马尼亚语的“pisic ̆a ro ̦sie”，就会遇到一些困难。后者是名词 + 形容词短语，因为罗马尼亚语中这类短语的自然语序与日语语序相反。

罗马尼亚语中的归约规则为

$$n_R(n_R^r n_R) \to n_R.$$

假设存在一个幺半函子 $T' : J \to R$，它将日语语法类型映射到罗马尼亚语语法类型。那么我们希望保留归约规则，即

$$T'\left((n_J n_J^\ell) n_J\right) = n_R(n_R^r n_R).$$

因此我们得到条件：$T'\left(n_J^\ell\right) = n_R^r$。然而，强幺半函子必须保留左伴随与右伴随。因此，这个条件无法满足。

第4 节介绍了有助于我们克服此类语序变化问题的技巧。

3. Japanese crash course

3.1. 概述。日语是一种综合性黏着语。常见语序是主语-宾语-动词（SOV），采用主题-述题句式结构。日语没有定冠词/不定冠词，名词也没有语法性别或数的变化。动词和形容词会根据时态、语态和体进行变形，但不随人称或数变化。助词附加在词语后，用于标示其语法功能。我们以日语原文书写，并采用日本式罗马字转写系统。

句子“猫吃鱼”可以用两种不同但密切相关的方式表示。

(1) 猫 が 魚 を 食べる neko ga sakana wo taberu 猫 NOM鱼 ACC 吃 (2) 猫 は 魚 を 食べる neko ha sakana wo taberu 猫 TOP鱼 ACC 吃

注意主语助词“ga”、主题助词“ha”和直接宾语助词“wo”的用法。请注意，日语区分主题和主语。主题通常需要在话语开头明确引入，但随着话语展开，主题不必是每个句子的语法主语。这两个句子都可以译成英语“The cat eats fish.”（猫吃鱼。）或“Cats

eat fish.”（猫吃鱼。）不过，第二个句子更贴切的理解是“至于猫／说到猫，它吃鱼。”

日语语序的另一个重要方面是中心语后置。短语大致可以描述为由中心语和修饰语组成。英语通常是一种中心语前置语言。例如，考虑以下短语：“to school”（去学校）、“in England”（在英格兰）和“red cat”（红色的猫）。被修饰的词通常位于修饰语之前，主要的例外是名词位于修饰它们的形容词之后。相比之下，日语可谓是一种典型的中心语后置语言。我们的示例短语变为

(3) 学校 へへへ
gakk¯o he 学校 去
(4) イギリスににに
igirisu ni 英格兰
在 (5) 赤い 猫猫猫
akai neko 红色 猫

中心语后置现象也见于关系从句：关系从句通常位于它们所修饰的词类之前。下面这组短语展示了这一现象。

(6) 女 が 赤い ワンピース を 着た onna ga akai
wanpˆısu wo kita 女人 NOM 红色 连衣裙 ACC
穿着 ‘女人穿了一条红色连衣裙’ (7) 赤い ワンピース
を 着た 女女女 akai wanpˆısu wo kita onna 红色 连
衣裙 ACC穿着 女人 ‘穿着红色连衣裙的女人’

这是一个典型示例，展示了翻译过程中语序发生变化的结构。下一节将介绍用于理解此类翻译的代数工具。

主语在语境明确时通常会被省略，人称代词也很少使用。本节最后以一个示例作结，说明一个十分常见的反身代词/人称代词“zibun”（“自己”）如何导致歧义解读。“Zibun”常被说话者用来指代自己或谈话对象。句子

(8) 自分 が 嘘つき か zibun ga usotuki ka
自己 NOM骗子 QUESTION

在缺乏更多语境的情况下，可以译为“我是骗子吗？”或“你是骗子吗？”。

3.2. 组合模型。 定义 $J = \mathrm{PGrp}(\{\pi, n, s_1, s_2, s, o_1, \ldots\})$ 为与日语相关的语法类型的预群。稍作修改后，我们遵循（[Car02]），定义以下原子类型：

- π代词，- n名词，- s_1, s_2 未完成体/完成体句子，- $\overline{s}$话题化句子，- s句子，- o_1 主格，- o_2 宾格，- o_3 与格，- o_4 属格，- o_5 处所格，- o_6 趋向格，- o_7 离格，- 等等。

我们还在 J 中施加以下约简：

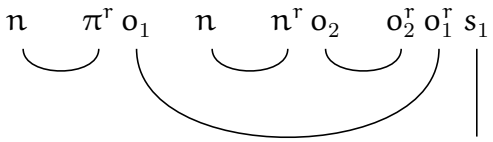
$$s_i \to s \qquad \overline{s} \to s \qquad n \to \pi_\circ$$

现在我们来讨论如何为各种词类指定类型。回顾示例句子“neko ga sakana wo taberu”（“猫吃鱼”），“neko”和“sakana”都是名词，因此类型为 n。主语助词“ga”的类型为 $\pi^r o_1$，直接宾语助词“wo”的类型为 $n^r o_2$，而及物动词“taberu”的类型则为 $o_2^r o_1^r s_{1\circ}$ 因此，该句子的类型为 $n(\pi^r o_1)n(n^r o_2)(o_2^r o_1^r s_1)$，我们可以推导出以下类型约简：

$$\begin{aligned} n(\pi^r o_1)n(n^r o_2)(o_2^r o_1^r s_1) &\to (n\pi^r)o_1(\pi\pi^r)o_2(o_2^r o_1^r s_1) \\ &\to (\pi\pi^r)o_1(\pi\pi^r)o_2(o_2^r o_1^r s_1) \\ &\to o_1 o_2(o_2^r o_1^r s_1) \\ &\to o_1(o_2 o_2^r)o_1^r s_1 \\ &\to o_1 o_1^r s_1 \\ &\to s_1 \\ &\to s \end{aligned}$$

从而可以看出该句子结构良好，并且约简为正确的语法类型。这里我们使用了约简 $n \to \pi$ 和 $s_1 \to s$，并以不同方式应用了收缩态射 $\varepsilon_\circ$ 从图示上看，这种类型约简可见于

下图，其中下方的括号表示应用了类型为 ε的收缩态射。



由于语序灵活，同一个句子也可以写成“sakana wo neko ga taberu”，这时 taberu 就会被赋予类型$o_1^r o_2^r s_1$。为了在进行计算时利用切换引理，我们希望将工作范围限制在自由生成的预群内。通常，对于某些词或动词可以被赋予不同类型的情况，可以通过添加元规则来处理。非正式地说，元规则规定：如果某个语法包含符合指定模式的规则，那么它也包含符合另一指定模式的规则。在这个具体例子中，我们可以施加以下元规则。

元规则 1. 任何类型为 $o_1^r o_2^r s_i$ 的及物动词，类型也为 $o_2^r o_1^r s_i$。

从及物动词转向其他内容，离格助词“kara”的类型为 $\pi^r o_7$，而向格助词“he”的类型为 $\pi^r o_6$。在下面的例子中，动词“un- tensita”的类型为 $o_6^r o_7^r s_2$。

(9) 家 从 车站 驾车了 ie kara eki he
untensita 房子 ABL 车站 LAT 开了车‘（我）从家开车去了火车站。’

使役被动动词有一个主语和一个由类型为 $\pi^r o_3$ 的与格助词“ni”标记的间接宾语。例如，动词“yomaseta”（“x 让 y 读”）的类型为 $o_2^r o_3^r o_1^r s_2$。

(10) 先生 が 私 に 本 を 読ませた sensei ga watasi ni hon wo
yomaseta 老师 NOM 我 DAT 书 ACC 读-CAUSE-PAS‘老师让我读这本书。’

属格助词“no”的类型是 $\pi^r o_4$，并附有一条元规则，指出类型 o_4 等同于类型 nn^ℓ。在属格结构中，领属者总是在左侧。主题助词 “ha” 与主语助词 “ga” 有所区别，其类型为 $\pi^r \bar{s} s^\ell$，也就是说，“ha”要求左侧是主题，右侧是关于该主题的句子。

(11)	私	の	車	は	箸	を	渡れない
	watasi	no	kuruma	ha	hasi	wo	watarenai
	I	GEN	car	TOP	bridge	ACC	cross-POT-NEG

“我不能开车穿过这座桥／关于我的车，它不能穿过这座桥。”

在后一示例中，类型约简过程如下：

$$\pi(\pi^r o_4)n(\pi^r \overline{s} s^\ell)n(\pi^r o_2)(o_2^r s_1)$$

$$\begin{aligned}
&\to(\pi\pi^r)o_4n(\pi^r\overline{s}s^\ell)(n\pi^r)(o_2o_2^r)s && \text{结合律}\\
&\to(1)(nn^\ell)n(\pi^r\overline{s}s^\ell)(n\pi^r)(1)s && \text{收缩 } +\text{属格元规则}\\
&\to n(n^\ell n)(\pi^r\overline{s}s^\ell)(\pi\pi^r)s && n\to\pi\\
&\to(n\pi^r)\overline{s}(s^\ell s) && \text{结合律}\\
&\to\overline{s} && \text{收缩}\to s_\circ
\end{aligned}$$

4. Translation and decorated pregroups

4.1. 装饰预群。 如示例 2 所示，我们最初的工具并不适合在语序不同的语言之间翻译短语。预群的态射（或幺函子） $T : P \to Q$ 用于将信息从源语言传递到目标语言，但它过于僵化。我们为预群添加额外结构，以便更好地控制幺半群的运算。为此，我们定义反同态来反转语序，并引入带花括号的预群以及 β-预群，以便更精细地控制结合律。

定义 6。幺半群反同态是一个映射 $\Phi : P \to Q$，满足对所有元素 x、$y \in P$，都有 $\Phi(xy) = \Phi(y)\Phi(x)$。

定义 7。设 $(P, \cdot)$ 是一个幺半群。对偶幺半群 $(P^{op}, *)$ 是一个幺半群，其元素与 P 相同，且对所有 x、$y \in P^{op}$，运算定义为 $x * y = y \cdot x$。显然可以观察到 $(P, \cdot) \simeq (P^{op}, *)$。

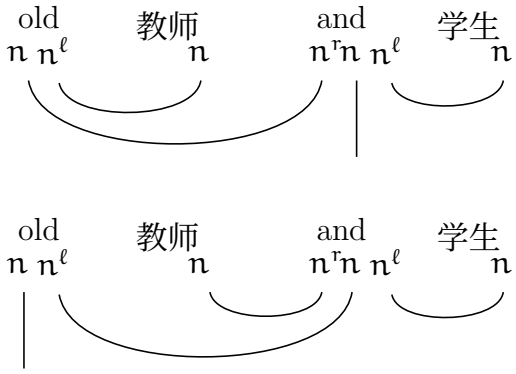
鉴于此，反同态可以视为从对偶幺半群 $\Phi : P^{op} \to Q$ 出发的态射。此外，预群的反同态会将左伴随映为右伴随，反之亦然。

例3。 在示例2 中，将日语中的“形容词 + 名词”短语翻译成罗马尼亚语的问题，可以通过将翻译函子设为一个反同态来解决，该反同态将 $n_J \mapsto n_R$映射到。然后，函子 T 保留所需的约化

$$T((n_J n_J^\ell)n_J) = T(n_J)\left(T(n_J^\ell)T(n_J)\right) = n_R\,(n_R^r n_R) \to n_R.$$

对较长短语和完整句子进行句法分析，会带来新的复杂性。例如，在简单的短语中，通常只有一种方式可以进行类型归约，从而确定短语的句法类型。结合律可能会在短语句法分析中引入歧义。下面的示例对此进行了说明。

例4。考虑短语“年长的教师和学生”。我们给“教师”和“学生”赋予类型 n。我们给形容词“年长的”赋予类型 nn^{ℓ} 。该短语中的连词“和”需要两个名词类型的输入，以生成一个名词短语，因此赋予它 $n^{r}nn^{\ell}$。我们可以利用幺半群运算的结合律，进行两种不同的类型归约。



两种类型归约都得到了所需的名词短语。不过，这两种解读略有不同。第一种将形容词“old”仅修饰“teachers”，因此句子被解析为“（年长的教师）和学生”；而第二种类型归约则认为“old”同时修饰“teachers”和“students”，形成短语“年长的（教师和学生）”。

我们可以构造这样的示例：改变约简的顺序，可能导致最终约简为合乎语法的句子，也可能导致约简为语法上无法接受的短语。因此，可以在预群中加入模态或 β- 结构，以局部抑制结合律。这样可以确保短语约简为正确的类型，或确保修饰语按照规定的方式分配。

带模态的预群最初由 [Fad02] 提出，其逻辑随后在 [KM07]中得到了更深入的研究。

定义 8. β-预群是一个预群 $(P, \cdot, 1, -^{\ell}, -^{r}, \leqslant)$，并附带一个单调映射 $\beta : P \to P$，使得 β 有一个右伴随 $\beta : P \to P$；也就是说，对所有的 x、$y \in P$，当且仅当 $x \leqslant \beta(y)$时，有 $\beta(x) \leqslant y$。

在实践中，我们会为预群语法中的类型添加模态，以表明某些约简必须优先进行。

例5。在前一个示例中，我们可以通过指定类型来规定句法分析“（年长的教师）和学生”

$$n\,[\beta(n)]^{\ell} \cdot [\beta(n)] \cdot n^{r}nn^{\ell} \cdot n$$

并通过指定类型来进行“老（教师和学生）”的句法分析

$$nn^{\ell} \cdot [\beta(n)] \cdot [\beta(n)]^{r}\, nn^{\ell} \cdot n.$$

现在，我们有办法在全局范围内反转语序，并在局部范围内阻止结合律。我们以介绍一种新的装饰类型来结束本节，这种装饰可以让我们在局部控制语序。

还建议读者参阅 [Sta08]，了解元组预群的入门介绍；参阅 [Lam10]，了解使用预群的积分析法语句子；并参阅 [Bob23]，了解具有局部预循环性的预群。

接下来，我们引入一种新的预群装饰。

定义 9. 带 k-括号的幺半群 $(P, \cdot, 1)$ 是一个自由幺半群，其中每个词都是指定子词的规定拼接。将此定义及后续定义推广到带 k-括号的预群是显而易见的。

示例 6. 考虑由两个字母生成的自由幺半群 $F = \mathrm{Mon}(\{a, b\})$。将 F 视为带 2-括号的幺半群时，$\langle abba\rangle\langle b\rangle$和 $\langle abb\rangle\langle ab\rangle$ 是不同的词，因为它们的指定子词不同。

定义 10. 带 k-括号的幺半群的态射 $f : (P, \cdot) \to (Q, *)$ 是保持指定子词的幺半群态射 $f : (P, \cdot) \to (Q, *)$。用符号表示：

$$f(\langle w_1\rangle \cdot \ldots \cdot \langle w_k\rangle) = \langle f(w_1)\rangle * \langle f(w_2)\rangle * \ldots * \langle f(w_k)\rangle.$$

例7（一些有用的构造）。我们定义两个带括号的幺半群态射，它们有助于理解翻译。

首先，设 P 是一个带 2-括号的幺半群，并考虑一个词

$$w = \langle w_1\rangle\langle w_2\rangle.$$

由于 P 的底层幺半群是自由的，我们可以将 w 视为自由积 $P * P \simeq P$ 的一个元素，其中指定子词 w_i 属于第 i 个因子。取以下幺半群态射序列

$$\Psi : P \simeq P * P \xrightarrow{f} P \times P \xrightarrow{g} P^{op} \times P^{op} \xrightarrow{h}$$

$$\xrightarrow{h} P^{op} * P^{op} \xrightarrow{i} Q$$

这里，f 是典范满射，发送 $\langle w_1\rangle\langle w_2\rangle \mapsto (w_1, w_2)$；$g$ 是一对反同构，它们在原子类型 $(w_1, w_2) \mapsto (w_1^{op}, w_2^{op})$ 上的作用如同恒等映射；h 是典范单射，发送 $(w_1^{op}, w_2^{op}) \mapsto \langle w_1^{op}\rangle\langle w_2^{op}\rangle$；而 i 是某个固定的带 2-括号幺半群的同态。

其次，令 P 为一个带 3-括号的幺半群。我们以类似的方式构造以下态射。

$$\Xi : P \simeq P * P * P \longrightarrow P \times P \times P \longrightarrow P \times P^{op} \times P \longrightarrow$$

$$\longrightarrow P * P^{op} * P \longrightarrow P * P * P \simeq P \longrightarrow Q$$

现在我们通过具体示例来说明短语翻译。本节余下部分将使用两个函子语言模型：用于日语的 $\mathcal{J}\colon \mathsf{J} \to \mathbf{FVect}$ 和用于英语的 $\mathcal{E}\colon \mathsf{E} \to \mathbf{FVect}$。我们还采用以下实用元规则。

元规则2。任何类型为 $so_1^r w$ 的动词也具有类型 $o_1^\ell s w$，其中 w 表示其余所有必需的补足语。

4.2. “有”／“存在”。日语有两个表示存在的动词“いる”和“ある”，分别用于有生命和无生命的存在。它们大致都表示“是”，不过更常见的英语译法是“有／存在”。

请看下面这个句子。

(12) 森 に 猫 が いる mori ni neko ga iru 森林 LOC 猫 NOM存在

人工译者可以用许多方式处理这个句子。一种标准且逐字的主谓宾翻译是“森林里有一只猫。”一种更自然的主谓宾译法是“有一只猫住在森林里。”考虑到这是写给儿童的短篇故事，甚至可以选择“森林里住着一只猫”，为文本营造童话般的氛围。本文选择采用直接的反同态进行翻译，因此目标译文是“森林里有一只猫。”

我们采用以下语法约化模型：

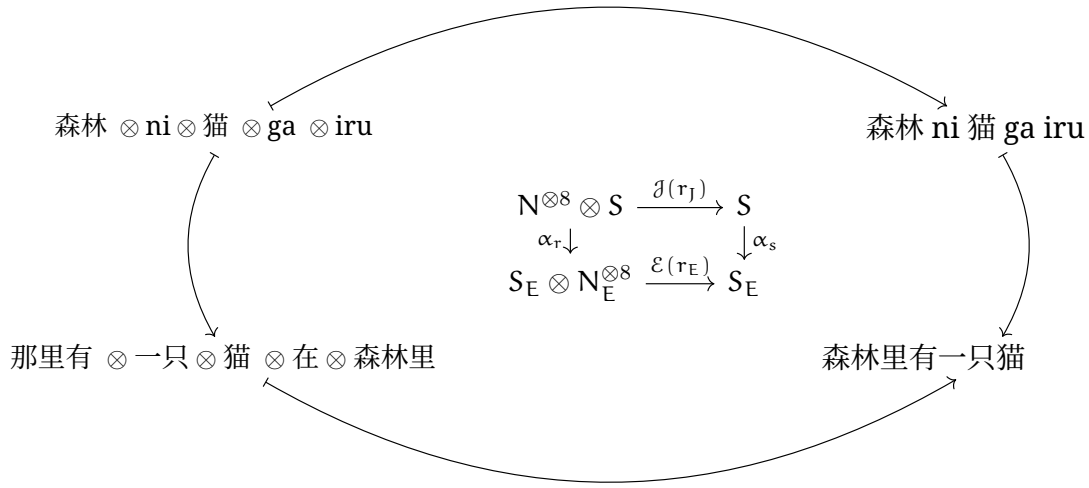
$$\mathsf{J} = \mathrm{PGrp}(\{n, o_1, o_5, s\}) \text{ and } \mathsf{E} = \mathrm{PGrp}(\{n_E, o_{1E}, o_{5E}, s_E\}).$$

在句法层面，翻译函子由反同态$\mathsf{T} : \mathsf{J} \to \mathsf{E}$ 给出，它将 $n \mapsto n_E, o_1 \mapsto o_{1E}$, $o_5 \mapsto o_{5E}, s \mapsto s_E$映射为。在语义层面，我们有 $F(n) = F(o_1) = F(o_5) = N$, $F(s) = S$ 和 $G(n_E) = G(o_{1E}) = G(o_{5E}) = N_E, G(s_E) = S_E$。

在 J中，我们有类型约简 $r = n(n^r o_5)n(n^r o_1)(o_1^r o_5^r s) \leqslant s$。应用翻译函子 T 后，我们得到：

$$\begin{aligned}
T(n(n^r o_5)n(n^r o_1)(o_1^r o_5^r s)) &= T(s)T(o_5^r)T(o_1^r)T(o_1)T(n^r)T(n)T(o_5)T(n^r)T(n) \\
&= (s_E o_{5E}^\ell o_{1E}^\ell)o_{1E}(n_E^\ell n_E)o_{5E}(n_E^\ell n_E) \\
&\to s_E o_{5E}^\ell (o_{1E}^\ell o_{1E})o_{5E} \\
&\to s_E (o_{5E}^\ell o_{5E}) \\
&\to s_E .
\end{aligned}$$

在语义层面，我们定义自然变换 $\alpha : \mathcal{J} \Rightarrow \mathcal{E} \circ T$按预期方式发挥作用，即映射 $N \to N_E$ 将 neko 映射为猫 $\mapsto$，将 mori 映射为森林 $\mapsto$；映射 $S \to S_E$ 将 iru 映射为「在那里」 $\mapsto$。我们还规定 ga 表示「一个」 $\mapsto$，ni 表示「在.....里」 $\mapsto$。下面图的交换性显而易见。



4.3. 简单的主宾谓句。我们将介绍一种翻译以下句子的方法。

(13) 医者 は 手紙 を 書く issya ga
tegami wo kaku 医生 NOM 信 ACC
写 '医生写了一封信。'

我们使用以下语法

$$J = \mathrm{PGrp}(\{n, o_1, o_2, s\}) \text{ and } E = \mathrm{PGrp}(\{n_E, o_{1E}, o_{2E}, s_E\}).$$

"issya"和"tegami"这两个词被赋予名词类型 n，"ga"和"wo"这两个助词分别具有通常的类型 $n^r o_1$ 和 $n^r o_2$，而及物动词"kaku"的类型为 $o_2^r o_1^r s$。这个句子显然是合乎语法的：$n(n^r o_1)n(n^r o_2)(o_2^r o_1^r s) \to s$。

这里我们采用带二重括号的预群这一概念。原则上，对于主宾谓句，我们按如下方式赋予括号：$\langle S\rangle\langle OV\rangle$。在我们的这个特定句子中，这变为

$$\big\langle n(n^r o_1)\big\rangle\big\langle n(n^r o_2)(o_2^r o_1^r s)\big\rangle.$$

我们将翻译函子定义为 Ψ 在示例7中定义的带括号的幺半群态射。结合元规则2，得到：

$$\begin{aligned}\Psi\big\langle n(n^r o_1)\big\rangle\big\langle n(n^r o_2)(o_2^r o_1^r s)\big\rangle &= \big\langle (o_{1E} n_E^\ell) n_E\big\rangle\big\langle (s_E o_{1E}^\ell o_{2E}^\ell)(o_{2E} n_E^\ell) n\big\rangle \\ &= \big\langle (o_{1E} n_E^\ell) n_E\big\rangle\big\langle (o_{1E}^r s_E o_{2E}^\ell)(o_{2E} n_E^\ell) n\big\rangle\end{aligned}$$

接下来，α 可以在原子类型上定义如下：issya $\mapsto$ 医生，tegami $\mapsto$信，kaku $\mapsto$ 写，而翻译 (Ψ, α) 给出

"(A/The) doctor write(s) (a/the) letter."

同样，冠词以及“write”变为第三人称单数的动词变位，既可以通过为助词“ga”和“wo”添加意义，以蛮力方式纳入我们的模型；也可以在翻译过程的另一个步骤中，单独检查一致关系和冠词。

4.4. 关系从句。 用预群来解释各种语言中的关系代词，颇具挑战性。在 [SCC13] 和 [SCC14], 中，作者在预群上添加了弗罗贝尼乌斯代数这一额外结构。通俗来说，弗罗贝尼乌斯代数结构为 ε， η 函子化的运用增添了额外的映射，其中最重要的称为“复制映射”和“反复制映射”。这些新的态射有助于更好地追踪短语内部的信息。例如，在英语句子

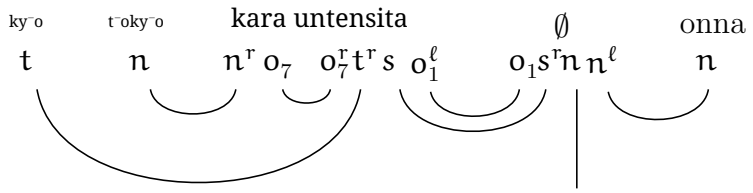
“那位今天从东京开车来的女士，参加派对迟到了”

新的态射可以形式化这样一个事实：主句“那位女士参加派对迟到了”的主语与关系从句“今天从东京开车来的”的主语是同一个人。关系代词“who”充当一座桥，将主语“复制”到关系从句中，然后再将其传回主句。

我们来翻译下面这个关系从句。

(14) 今日 東京 から 運転した 女 ky⁻o t⁻oky⁻o
kara untensita onna 今天 东京 ABL 开车了 女人 ‘今天从东京开车来的女人。’

我们以一种不那么直截了当的方式赋予类型。首先，我们在修饰语“t⁻oky⁻o kara untensita”和中心语“onna”之间插入一个空词。我们赋予以下类型：“t⁻oky⁻o”和“onna”的类型都是 n，离格助词“kara”的类型是 $n^r o_7$，“ky⁻o”的类型是 t（时间副词），动词“untensita”的类型是$o_7^r t^r s o_1^\ell$，而这个空词的作用类似于一个类型为 $o_1 s^r n n^\ell$的虚拟关系代词。

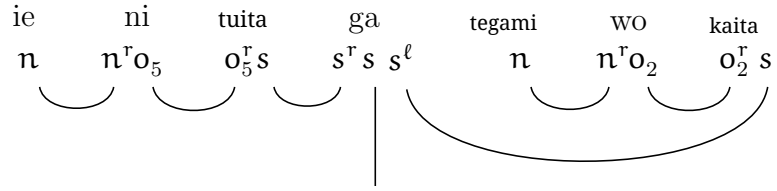


这一构造生成一个名词短语，可以用一个简单直接的反同态来翻译。这种巧妙构造的优点在于，现在我们可以把空词翻译成关系代词“who”或“that”。这与 [SCC13]的弗罗贝尼乌斯代数方法完美契合。在这个示例中，我们将关系从句建模为参考文献作者所称的主语关系从句。

4.5. 并列句。连接句子以构成并列句，最简单的方法是用助词“ga”（“和”）将它们连接起来，并为其指定类型 $s^r s s^\ell$。我们将翻译下面这个省略了主语的句子。

(15) 家 に 着いた が 手紙 を 書いた ie ni tuita ga tegami wo kaita 房子 LOC到达了，并且信 ACC 写了 ‘我到家了，并写了一封信。’

在日语中，我们有如下约简图。



我们用花括号装饰预群，并赋予以下类型

$$\Big\langle n \cdot n^r o_5 \cdot o_5^r s \Big\rangle \Big\langle s^r s s^\ell \Big\rangle \Big\langle n \cdot n^r o_2 \cdot o_2^r s \Big\rangle.$$

将态射 Ψ 从示例 7 扩展到带有 3-花括号的幺半群，我们得到 $\Psi\langle n\cdot n^r o_5\cdot o_5^r s\rangle\langle s^r s s^\ell\rangle\langle n\cdot n^r o_2\cdot o_2^r s\rangle = \langle s_E o_{5E}^\ell \cdot o_{5E} n_E^\ell \cdot n_E\rangle\langle s_E^r s_E s_E^\ell\rangle\langle s_E o_{2E}^\ell \cdot o_{2E} n_E^\ell \cdot n_E\rangle$。

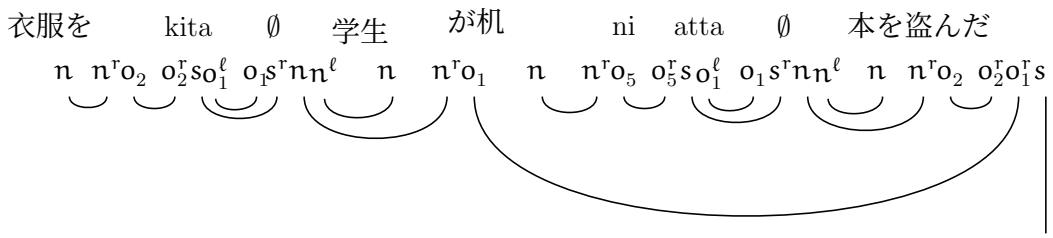
假设省略的主语指的是第一人称单数，对其应用一个适当定义的 α 后，翻译为

“(I) arrived home and (I) wrote (a) letter.”

4.6. 综合运用。我们结合所有技巧来研究一个更复杂的句子。

(16) 穿着 制服 的 学生 偷了 桌子 上 的 书 seihuku o kita gakusei ga tukue ni atta hon wo 制服 ACC 穿着 学生 NOM 桌子 LOC 上有 书 ACC盗んだ nusunda 偷了「那位穿着制服的学生偷了桌上的书。」

这是一个标准主宾谓句，其中主语和直接宾语都由关系从句修饰。在日语预群语法中，我们可以进行以下直接的化简。



可以看到，在上图中，我们利用结合律证明该句子能够化简为正确的句法类型。为了确保化简和翻译万无一失，我们用花括号和 a β-结构对预群语法进行标注。随后为该句子赋予类型

$$\langle \mathrm{n}\cdot\mathrm{n}^{\mathrm{r}}\mathrm{o}_2\cdot\mathrm{o}_2^{\mathrm{r}}\mathrm{o}_1^{\ell}\cdot\mathrm{o}_1\mathrm{s}^{\mathrm{r}}\mathrm{n}\boldsymbol{\beta}(\mathrm{n}^{\ell})\cdot\boldsymbol{\beta}(\mathrm{n})\cdot\mathrm{n}^{\mathrm{r}}\mathrm{o}_1\rangle\langle \mathrm{n}\cdot\mathrm{n}^{\mathrm{r}}\mathrm{o}_5\cdot\mathrm{o}_5^{\mathrm{r}}\mathrm{so}_1^{\ell}\cdot\mathrm{o}_1\mathrm{s}^{\mathrm{r}}\mathrm{n}\boldsymbol{\beta}(\mathrm{n}^{\ell})\cdot\boldsymbol{\beta}(\mathrm{n})\cdot\mathrm{n}^{\mathrm{r}}\mathrm{o}_2\cdot\mathrm{o}_2^{\mathrm{r}}\mathrm{o}_1^{\mathrm{r}}\mathrm{s}\rangle$$

并应用 Ψ 示例 7 中的态射，再结合元规则 2，该句子翻译为

“（某个/这个）穿着（某套/这套）制服的学生偷了（一本/这本）书，而这本书是（LOC）桌子。”

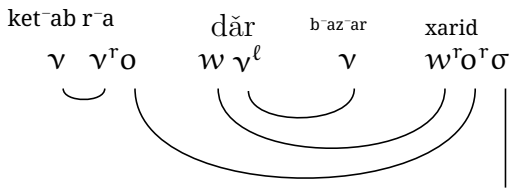
4.7. 从波斯语到日语的一个示例。波斯语与日语有某些相似之处，因此翻译（至少在句法层面上）会相对简单一些。例如，波斯语也采用主宾谓语序，名词没有语法性别，而且是一种代词省略语言。一个重要的结构差异是，波斯语既使用前置词，也使用后置词作为格标记。

以下 [Sad07]，我们使用以下（简化的）预群来模拟波斯语语法 $\mathsf{F} = \mathrm{PGrp}(\{\nu, \sigma, \mathrm{o}, w\})$，其中原子类型分别表示名词、句子、直接宾语和介词短语。在日语一侧，我们使用$\mathsf{J} = \mathrm{PGrp}(\{\mathrm{n}, \mathrm{s}, \mathrm{o}_2, \mathrm{o}_5\})$，其含义与通常相同。将两个函子语言模型分别记为 $\mathcal{F} : \mathsf{F} \to \mathbf{FVect}$ 和 $\mathcal{J} : \mathsf{J} \to \mathbf{FVect}$。

我们想把下面这个句子从波斯语翻译成日语。

(17) ket ⁻ab r ⁻a dˇar b ⁻az ⁻ar xarid book
ACC PREP market bought ‘他/她从市场买了一本书。’

这里，“ket⁻ab r⁻a”是直接宾语，“dˇar b⁻az⁻ar”是介词短语，而“xarid”是过去时的及物动词。这个示例句省略了主语，并同时使用后置词“r⁻a”和介词“dˇar”来标记格。在波斯语中，我们有如下归约。



函子语言模型 $\mathcal{F}, \mathcal{J}$ 发送 $\nu, o, w \mapsto N_F$（波斯语名词）和 $\sigma \mapsto S_F$（波斯语句子），以及 $n, o_2, o_5 \mapsto N_J$（日语名词）和 $s \mapsto S_J$（日语句子）。自然变换 $\alpha : \mathcal{F} \Rightarrow \mathcal{J} \circ \underline{i}$ 由 ketab $\mapsto$ ho$\underline{n}$、ra $\mapsto$ o、d$\breve{}$a $\mapsto$ de、bazar $\mapsto$ itiba 和 xarid $\mapsto$ kaimasita 给出。在句法层面，我们定义了一个幺半翻译函子 $T : F \to J$，它映射 $\nu \mapsto n, \sigma \to s, o \to o_2$ 和 $w \to o_5$。

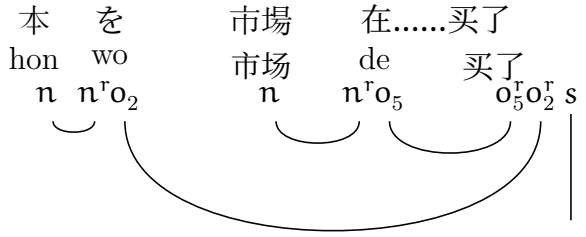
预群用 3-花括号修饰。句子被赋予类型

$$\langle \nu \cdot \nu^r o \rangle \langle w \nu^\ell \cdot \nu \rangle \langle w^r o^r \sigma \rangle.$$

从句法上看，翻译函子取为 Ξ，来自示例 7。语序如下调整

$$\Xi \langle \nu \cdot \nu^r o \rangle \langle w \nu^\ell \cdot \nu \rangle \langle w^r o^r \sigma \rangle = \langle n \cdot n^r o_2 \rangle \langle n \cdot n^r o_5 \rangle \langle o_5^r o_2^r s \rangle$$

这就得出了日语中的以下类型约简。



5. Future work

本文介绍了装饰预群，并将其用作构建不同语序的自然语言之间组合式翻译概念的一种手段。我们的目标是证明，在不完全牺牲函子性的情况下，仍可采用范畴论方法对翻译进行建模。我们的一些构造是临时设计的，其中大多数仍有改进空间。

首先，需要解决的是如何将名词没有语法性别和数的语言翻译成具有这些特征的语言。使用积预群或元组预群来处理语法一致，或许是一种可行的方向，尽管要找到实现这一点的直接模型似乎并不容易。

其次，我们可以研究特征差异和结构差异更大的语言之间的翻译。例如，我们该如何（函子性地）解释

具有主宾格对齐的语言与具有作格—绝对格对齐（或分裂作格对齐）的语言之间的翻译?

第三，本文重点关注句法。探索如何在不同说话者之间协商翻译中的意义，以及如何追踪他们不断变化的语义空间，将会很有意思。从更技术的角度来说，可以将意义空间从 FVect 换成具有更丰富结构的范畴，例如 ConvexRel。该范畴的对象是凸代数，态射则是凸关系。在 [BCG+19] 中，作者证明了 ConvexRel 是一个紧闭对称幺半范畴，因此适用于在组合分布式函子语言模型中对语义进行建模。

最后，撇开翻译问题不谈，还可以关注如何扩展 Cardinal（[Car02] [Car06] [Car07]）的研究，并提出一种更完善的预群方法，用以分析日语语法中其他典型的体。作者尤其感兴趣的是并列句、从句以及内中心关系从句的结构。

References

[BCG+19]Joe Bolt、Bob Coecke、Fabrizio Genovese、Martha Lewis、Dan Marsden 和 Robin Piedeleu。《交互概念空间 I：概念的语法组合》，2019年。DOI:10.1007/978-3-030-12800-5 9.[BL01a] Daniele Bargelli 和 Joachim Lambek。《法语句子结构的代数方法》，2001年。DOI:10.1007/3-540-48199-0 4.[BL01b] Donna Bargelli 和 Joachim Lambek。《阿拉伯语句子结构的代数方法》，2001年。Corpus ID: 56575156.[BLMT 18]Tai-Danae Bradley、Martha Lewis、Jade Master 和 Brad Theilman。《翻译与演化：迈向 DisCoCat 中语言变化模型》，2018年。DOI:10.4204/EPTCS.283.4.[Bob23] Valentin Boboc。《增强型预群及其在语言学中的应用》，2023年。arXiv:2303.05160.[Car02] Kumi Cardinal。《日语语法的代数研究》，2002年。博士论文.[Car06] Kumi Cardinal。《类型语法与日语助词相遇》，2006年.[Car07] Kumi Cardinal。《日语使役结构的预群分析》，2007年.[CL05] Claudia Casadio 和 Jim Lambek。《拉丁语语法的计算代数方法》，2005年。DOI:10.1007/s11168-005-1286-0.[CSC10] Bob Coecke、Mehrnoosh Sadrzadeh 和 Stephen Clark。《组合分布式意义模型的数学基础》，2010年。DOI:10.48550/arXiv.1003.4394.[Fad02] Mario Fadda。《迈向灵活的预群语法》，2002年.[KL80] Gregory M Kelly 和 Miguel L Laplaza。《紧闭范畴的相干性》，1980年。DOI:10.1016/0022-4049(80)90101-2.[KM07] Aleksandra Ki´slak-Malinowska。《β前群的逻辑》，2007年。DOI:10.1007/s11225-007-9090-5.[Lam97] Joachim Lambek。《重访类型语法》，1997年。DOI:10.1007/3-540-48975-4 1.[Lam10] Joachim Lambek。《通过并行预群计算探究法语中的特征一致》，2010年。DOI:10.1007/s10849-009-9098-5.[LP04] Joachim Lambek 和 Anne Preller。《德语句子的代数方法》，2004年。HAL ID: irmm-00108541.[ML08] Jeff Mitchell 和 Mirella Lapata。《语义组合的向量模型》，2008年。https://aclanthology.org/P08-1028.pdf。

[萨德07] Mehrnoosh Sadrzadeh。《波斯语句子的预群分析》，2007年。https://www.cs.ox.ac.uk/files/2416/PersPreGroup.pdf。[SCC13] Mehrnoosh Sadrzadeh、Stephen Clark 和 Bob Coecke。《词义的弗罗贝尼乌斯结构 I：主语和宾语关系代词》，2013年。DOI:10.1093/logcom/ext044。[SCC14] Mehrnoosh Sadrzadeh、Stephen Clark 和 Bob Coecke。《词义的弗罗贝尼乌斯结构 II：所有格关系代词》，2014年。DOI:10.1093/logcom/exu027。[Sta08] Edward P Stabler。《元组预群语法》，2008年。

School of Mathematics, The University of Manchester, Alan Turing Building, Oxford Road, Manchester M13 9PL, United Kingdom

电子邮件地址：valentinboboc@icloud.com