

# 文法変分自己符号化器を用いた自然言語文の連続的な統語構造変換

折口希実 (指導教員: 小林一郎)

## 1 はじめに

深層生成モデルとは、変分自己符号化器 (Variational Autoencoder) [1] や敵対的生成ネットワーク (Generative Adversarial Network) [2] が代表例として挙げられる深層ニューラルネットワークで構成された生成モデルである。深層生成モデルでは入力するデータは画像や音響などの連続なデータが一般的であるが、化学式や数式などの離散なデータを扱い、生成するためには離散な構造を持ったデータを深層学習の枠組みで連続的な形へ変換する必要がある。

本研究では、深層学習の枠組みで、文脈自由文法によって表される自然言語の統語構造を連続的に変換することを試みる。

## 2 文法変分自己符号化器

Kusner ら [3] は、文法変分自己符号化器 (Grammar VAE) を提案した。このモデルは、分子を対象としており、分子の化学構造を文字列化する SMILES 記法を用いている。GrammarVAE の仕組みを図 1 に示す。

GrammarVAE は連続なデータを扱う VAE を応用し、文法の生成規則を用いることで離散なデータを扱えるようにしている。主な仕組みは同じであるが、潜在空間を連続的に滑らかにすることによってそれを可能にしている。

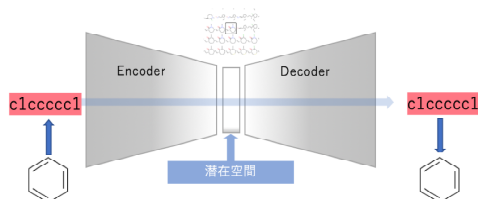


図 1: GrammarVAE の概要

## 3 統語構造を対象とした GrammarVAE

本研究では、自然言語の統語構造を対象としたモデルを生成した。その encoder と decoder の概要を図 2、図 3 に示す。また、文法に関しては自然言語の構文解析に広く応用されてきた文脈自由文法を用いている。

### 3.1 GrammarVAE の仕組み

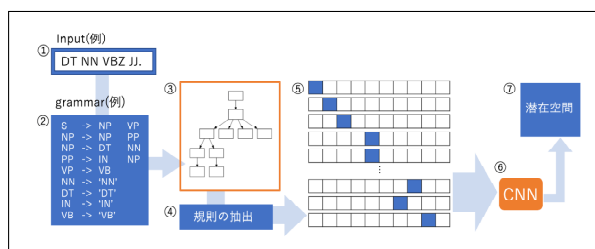


図 2: 自然言語の統語構造を適用した GrammarVAE の Encoder

**Encoding** 図 2 に示している流れが encoder の動きである。①にあるような自然言語の統語構造文を入力とする。また、②にあるような文法の生成規則を定義しておく。まず、入力した統語構造を文法に従って解析を行い、③に示すような木構造へ変換する。そしてその木構造から前順走査し、その解析で用いられた規則だけを抽出する。抽出した規則を⑤のような 1-hot ベクトルに変換する。このとき、抽出した規則の数だけベクトルは存在し、そのベクトルごとに全規則のうち抽出したベクトルが該当するようになっている。⑥では⑤で生成した複数の 1-hot ベクトルを 1 次元畳み込みニューラルネットワーク (Convolutional Neural Network) を用いて畳み込み、潜在空間へ落とし込む。

**Decoding** 図 3 に示している流れが decoder の動きである。まず潜在空間からベクトルを抽出する。再帰型ニューラルネットワーク (Recurrent Neural Network) を用いて③のように複数正規化されていないベクトル ('logits') を生成し、Last-in First-out (LIFO) スタックを用いて⑤に示すような生成規則をサンプリングする。サンプリングした生成規則の終端記号から⑥のような統語構造を生成する。

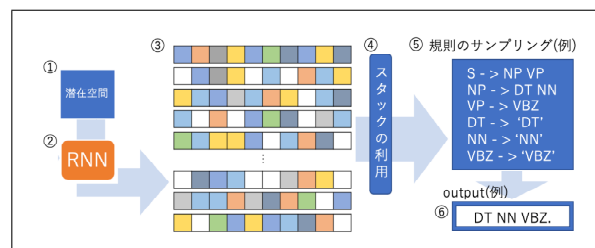


図 3: 自然言語の統語構造を適用した GrammarVAE の Decoder

**Training** Algorithm 1 に training アルゴリズムを示す。encoder での出力を  $X$ 、定義された生成規則の総数を  $K$  とする。decoder では、RNN の時間ステップ  $t$  の最大値を  $T_{max}$  とし生成された logits の集合は行列  $\mathbf{F} \in \mathbb{R}^{T_{max} \times K}$  と表せる。文法における非終端記号を  $\alpha$  とし、logits をマスクするベクトルを  $\mathbf{m}_\alpha \in [0, 1]^K$  としている。次の分布の式 (1) は時間ステップ  $t$  でマスクされた logits から生成規則を抽出するためのものである。また、 $f_{tk}$  は行列  $\mathbf{F}$  の  $(t, k)$  要素を指す。

$$p(\mathbf{x}_t = k | \alpha, \mathbf{z}) = \frac{m_{\alpha,k} \exp(f_{tk})}{\sum_{j=1}^K m_{\alpha,k} \exp(f_{tj})}. \quad (1)$$

### 3.2 連続的な統語構造変換

GrammarVAE において、潜在空間はガウス分布によって表されている。潜在空間の潜在変数を徐々に変化させることにより、連続的な統語構造変換することが可能である。

