

創発言語はなぜ Zipf 短縮に従わないのか？ 暗黙の事前分布に基づく再解釈

上田亮
東京大学

ryoryoueda@is.s.u-tokyo.ac.jp

概要

本稿の目的は、シグナリングゲームにおいて生じる創発言語が Zipf 短縮に従わないという知見に着目し、その原因の一端を解明することにある。これまで、創発言語が Zipf 短縮に従わないのは、エージェントの帰納バイアスせいだろうと考えられてきた。それに対し、本稿では新たな解釈を提示する。シグナリングゲームの目的関数の中で暗黙に仮定されている事前分布に着目し、それが反 Zipf 短縮の一因となっている可能性を示唆する。

1 はじめに

創発言語は、シミュレーション上においてエージェント同士でコミュニケーションを取らせるときに生じる、言語のようなプロトコルのことを指し、近年研究対象として注目されつつある [1]。創発言語の構造を理解し、それが自然言語とどの程度類似・乖離しているのかを知り、乖離をどのようにして埋めていけばよいのかを考えることが、当分野で重要なテーマの1つである。

Chaabouni ら [2] は、自然言語の普遍性質である Zipf 短縮 [3, 4, 5] を取り上げ、シグナリングゲーム [6] において創発する言語が、Zipf 短縮に従うか検証した。**Zipf 短縮**とは「頻出する単語ほど、単語長が短い傾向にある」という自然言語に普遍的にみられる性質である。例えば、英単語 “of” や “the” は高頻出であり、かつ短い。また、**シグナリングゲーム**はシンプルなコミュニケーションモデルであり、言語創発分野で頻繁に用いられる。シグナリングゲームには、**送信者・受信者**と呼ばれる2人のエージェントが登場し、送信者から受信者への単方向通信のみが許される。各プレイにおいて、送信者は**入力** $i \in I$ を1つ受け取り、**メッセージ** $m \in M$ を生成する。続いて、受信者はそのメッセージ m を受け取り、元

の入力が何であったのかを推測する。この推測が当たっていればゲームは成功である。このとき、送信者と受信者をつなぐメッセージ(の集合)は、何かしらの意味を伝達する言語のようなものであると考えることができる。[2] は、1メッセージがちょうど1単語を成すと見做せる設定の下、創発言語が Zipf 短縮に従うか検証した。ところが、標準的な設定のシグナリングゲームにおいて生じる創発言語は、Zipf 短縮に従わないことが示された。具体的には、メッセージが明らかに不必要に長くなってしまえばかりでなく、頻度の高いメッセージほど長くなるという、**反 Zipf 短縮**とでもいうべき傾向が示された。これまで、創発言語が Zipf 短縮に従わないのは、エージェントの元となるニューラルネットワーク (NN) アーキテクチャ (e.g., LSTM [7]) の帰納バイアスに問題があるからであろうと考えられてきた [2, 8, 9]。

それに対して、本稿では、反 Zipf 短縮が起こる原因について新たな解釈を提示する。シグナリングゲームの学習の際に用いられる目的関数の中で、暗黙に仮定されているメッセージの事前分布 $P(m)$ に着目し、その事前分布が反 Zipf 短縮の一因となっている可能性を示唆する。

2 背景

2.1 Zipf 短縮

Zipf 短縮 (Zipf's law of abbreviation) [3, 4, 5] は、言語学者 Gerge K. Zipf によって唱えられた自然言語の法則である。¹⁾²⁾ Zipf によれば、高頻度で出現する単語ほど長さが短くなる傾向にあるという。

2.2 シグナリングゲーム

言語創発では、シグナリングゲーム [6] (以降、単に**ゲーム**と呼ぶ) が頻繁に用いられる。ここでは

- 1) 所謂 Zipf 則 [10] とは異なる概念であることに注意。
- 2) 日本語訳は [5] に準拠。

特に [2] による設定を紹介する。ゲームには、有限の**入力集合** I 、有限の**メッセージ集合** M 、**送信者** $P_{\theta_s}(M = m | I = i)$ と **受信者** $P_{\theta_r}(I = i | M = m)$ が登場する。送・受信者エージェントは LSTM [7] 等の NN に基づいて定義され、各々学習可能なパラメータ θ_s, θ_r をもつ。 $I := \{1, 2, \dots, |I|\}$ とし、各 $i \in I$ は one-hot ベクトルで表現する。メッセージ集合 M は可変長メッセージの集合とする: $M := \{a_1 \dots a_{T-1} \text{eos} \mid a_i \in \mathcal{A} \setminus \{\text{eos}\}, 1 \leq i \leq T\}$ 。ただし、 \mathcal{A} は有限アルファベット、 $\text{eos} \in \mathcal{A}$ は EOS 記号、 T は最大長である。ゲームの流れは以下のようになる:

1. 入力 $i \in I$ をサンプリングする。ただし、 i は冪分布に従うものとする: $i \sim P_I(i) \propto i^{-1}$ 。
2. 送信者からメッセージ m をサンプリングする: $m \sim P_{\theta_s}(m | i)$ 。
3. 受信者による復元の程度を表す対数尤度 $\log P_{\theta_r}(i | m)$ を最大化することを目標とする。

$i \in I$ は one-hot ベクトルで表されるうえ、冪分布 $P_I(i)$ に従う (Zipf 則に従う) ので、一般的な NLP 分野でいうところの単語 ID とすることができる。故に各 i に対応するメッセージ m は単語と見做せる。

2.3 ゲームの目的関数

本稿では、最適化は勾配法で行うものとして議論を進める。ゲームの最適化は、以下の目的関数 $\mathcal{J}_{\text{game}}$ を勾配法で最大化することで行われる:

$$\mathcal{J}_{\text{game}} := \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}(i | m)]. \quad (1)$$

直観的には、受信者 R による復元の正確さを平均的に最大化するのが目的としていると理解できる。 $\mathcal{J}_{\text{game}}$ の微分は以下の形で得られる [2, 11]:³⁾

$$\begin{aligned} \nabla \mathcal{J}_{\text{game}} &= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\\ &\quad \nabla \log P_{\theta_r}(i | m) \\ &\quad + \{\log P_{\theta_r}(i | m) - b(i)\} \nabla \log P_{\theta_s}(m | i)]. \end{aligned} \quad (2)$$

ここで、 $b: I \rightarrow \mathbb{R}$ は任意のベースライン関数である。**式 2** は受信者を通常の誤差逆伝播によって、送信者を方策勾配 [12] によって最適化している。しかし、 $\mathcal{J}_{\text{game}}$ で創発する言語は Zipf 短縮とは真逆の傾向 (**反 Zipf 短縮**) に従うことが知られている。そこで、**メッセージ長に関する罰則項**を設け、メッセージが不必要に長くなるのを防ぐ場合がある:

$$\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha) := \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}(i | m) - \alpha |m|]. \quad (3)$$

3) ∇ は $\nabla_{\theta_s, \theta_r}$ の略記。以降でもこの略記を用いる。

ここで、 $\alpha \in \mathbb{R}_{\geq 0}$ は罰則項の強度を制御する超パラメータ、 $|m|$ はメッセージ m の長さを表す。 $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ の微分は以下である:

$$\begin{aligned} \nabla \mathcal{J}_{\text{game}}^{\text{LP}}(\alpha) &= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\\ &\quad \nabla \log P_{\theta_r}(i | m) \\ &\quad + \{\log P_{\theta_r}(i | m) - \alpha |m| - b(i)\} \nabla \log P_{\theta_s}(m | i)]. \end{aligned} \quad (4)$$

式 2 と比較すると、送信者の報酬が修正されたような形をしている。この罰則項により、Zipf 短縮が創発することが知られている [2, 8]。

3 ゲームの目的関数の分析

これまで、 $\mathcal{J}_{\text{game}}$ を最適化して得られる創発言語が Zipf 短縮に従わないのは、エージェントに適切な帰納バイアスがないからだと考えられてきた。例えば [8] は、送信者には laziness が欠けており、受信者には impatience が欠けていると主張している。しかし、エージェントのバイアスだけが原因ではないように思われる。本節では、目的関数 $\mathcal{J}_{\text{game}}$ の内に暗黙に仮定されている事前分布に着目し、それが反 Zipf 短縮の一因になっているのではないかと考える。以降の節にてそれを示唆する議論を展開する。

3.1 ゲームには暗黙の事前分布が存在する

ここで、やや唐突だが、**一様メッセージ分布** $P^{\text{uni}}(m) = |M|^{-1}$ を考え、受信者 $P_{\theta_r}(i | m)$ と合わせて、以下のような i, m の同時分布を定義する:

$$P_{\theta_r}^{\text{uni}}(i, m) := P_{\theta_r}(i | m) P^{\text{uni}}(m). \quad (5)$$

このとき、次の等式が成り立つ (証明は **付録 A** 参照):

$$\nabla \mathcal{J}_{\text{game}} = \nabla \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}^{\text{uni}}(i, m)]. \quad (6)$$

即ち、 $\mathcal{J}_{\text{game}}$ の最大化は、同時分布 $P_{\theta_r}^{\text{uni}}(i, m)$ の対数尤度の期待値を最大化するのと同じである。従って、 $\mathcal{J}_{\text{game}}$ という目的関数は $P^{\text{uni}}(m)$ をメッセージの**事前分布**として暗黙のうちに仮定していることになる。

次に、**モンキータイピング・メッセージ分布** $P^{\text{mon}}(m) = \frac{1}{Z} |\mathcal{A}|^{-|m|}$ (Z は適当な正則化項) を考え、受信者 $P_{\theta_r}(i | m)$ と合わせて、以下の i, m の同時分布を定義する:

$$P_{\theta_r}^{\text{mon}}(i, m) := P_{\theta_r}(i | m) P^{\text{mon}}(m). \quad (7)$$

このとき次の等式が成り立つ (証明は **付録 A** 参照):

$$\nabla \mathcal{J}_{\text{game}}^{\text{LP}}(\log |\mathcal{A}|) = \nabla \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}^{\text{mon}}(i, m)]. \quad (8)$$

即ち、 $\log P_{\theta_r}^{\text{mon}}(i, m)$ の期待値の最大化は、メッセージ長に関する罰則項の重みを $\alpha = \log |\mathcal{A}|$ としたときの $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ の最大化と同じである。更に一般化して、 $P^{\text{mon}}(m)$ と $P^{\text{uni}}(m)$ の中間的なメッセージ分布 $P^{\text{int}}(m; \eta)$ ($0 \leq \eta \leq |\mathcal{A}| - 1$) を考える:

$$P^{\text{int}}(m; \eta) = \frac{1}{Z'} (|\mathcal{A}| - \eta)^{-|m|}. \quad (9)$$

Z' は適当な正則化項である。 $P^{\text{int}}(m; \eta = 0) = P^{\text{mon}}(m)$ 及び $P^{\text{int}}(m; \eta = |\mathcal{A}| - 1) = P^{\text{uni}}(m)$ が成立するので、 P^{int} は中間的な分布といえる。これまでの議論と同様に、受信者 $P_{\theta_r}(i | m)$ と合わせて、 i, m の同時分布 $P_{\theta_r}^{\text{int}}(i, m; \eta) := P_{\theta_r}(i | m) P^{\text{int}}(m; \eta)$ を定義する。このとき、次式が成り立つ(証明は付録 A 参照):

$$\begin{aligned} \nabla \mathcal{J}_{\text{game}}^{\text{LP}}(\log\{|\mathcal{A}| - \eta\}) \\ = \nabla E_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}^{\text{int}}(i, m; \eta)]. \end{aligned} \quad (10)$$

即ち、 $\log P_{\theta_r}^{\text{int}}(i, m; \eta)$ の期待値の最大化は、罰則項の重みを $\alpha = \log\{|\mathcal{A}| - \eta\}$ としたときの $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ の最大化と同じである。逆に言えば、 $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ の最大化は、 $\eta = |\mathcal{A}| - \exp(\alpha)$ としたときの $\log P_{\theta_r}^{\text{int}}(i, m; \eta)$ の期待値の最大化と同じである。⁴⁾ $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ は、メッセージに関する事前分布 $P^{\text{int}}(m; \eta = |\mathcal{A}| - \exp(\alpha))$ を暗黙に仮定していることになる。なお、 $P^{\text{int}}(m; \eta = |\mathcal{A}| - 1) = P^{\text{mon}}(m)$ より、 $\mathcal{J}_{\text{game}}$ の暗黙の事前分布を $P^{\text{int}}(m; \eta = |\mathcal{A}| - 1)$ と言い直すこともできる。

3.2 暗黙の事前分布と反 Zipf 短縮

では、ゲームの目的関数の暗黙の事前分布はどのような形状をしているのだろうか。本節では事前分布の形状について議論し、そこから反 Zipf 短縮が生じる一因を垣間見る。事前分布 $P^{\text{uni}}(m)$, $P^{\text{mon}}(m)$, $P^{\text{int}}(m)$ はそのままでは可視化しづらいため、代わりに、対応するメッセージ長 $1 \leq l \leq T$ の分布を考える:

$$P^{\text{int}}(l; \eta) := \sum_{m \in \mathcal{M}} P^{\text{int}}(m; \eta) \mathbb{1}_{l=|m|}. \quad (11)$$

$P^{\text{uni}}(l)$, $P^{\text{mon}}(l)$ についても同様に定義する。⁵⁾ ここで、 $(T, |\mathcal{A}|) = (30, 4)$ としたときの分布の様子を図 1 に示す。図 1 から分かるように、 $P^{\text{uni}}(l)$ はかなり右に偏った形状をしている。つまり、 $P^{\text{uni}}(m)$ からメッセージ m をサンプリングしてきたとき、

4) η が $0 \leq \eta \leq |\mathcal{A}| - 1$ を満たさなくとも式 9 は確率分布になるが、先行研究 [2, 8] の α は必ず $\log |\mathcal{A}| \geq \alpha \geq 0$ に収まっているため、 η の範囲を限定しても特に問題はない。

5) ただし、自然言語の単語長の分布はガンマ分布に近いとされる [13]。逆に、 l の分布をガンマ分布と仮定してゲームの目的関数を逆算したらどうなるかについては付録 B 参照。

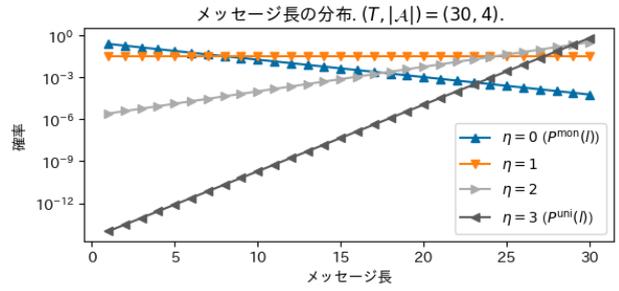


図 1 メッセージ長の分布 (($T, |\mathcal{A}|$) = (30, 4)).

m の長さは高い確率で最大長 T かその付近になる。ゲームの目的関数 $\mathcal{J}_{\text{game}}$ の暗黙の事前分布が P^{uni} であったことを踏まえると、 $\mathcal{J}_{\text{game}}$ の最適化で創発するメッセージが不必要に長くなってしまふのは、今や当然のことのように思える。一方、 $P^{\text{mon}}(l)$ は左に偏った分布になっていて、 $P^{\text{mon}}(m)$ それ自身が Zipf 短縮に従うメッセージ分布になっていることが分かる。⁶⁾ $P^{\text{int}}(l; \eta)$ は、 η の値によって $P^{\text{mon}}(l)$ と $P^{\text{uni}}(l)$ の間をつなぐような形状になっている。 $P^{\text{int}}(l; \eta)$ が、メッセージ長に関する罰則項を設けたゲームの目的関数 $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ における暗黙の事前分布になっていたことを踏まえると、罰則項の重み α を調整することに対応しているのだと理解することができる。

次に、 $(T, |\mathcal{A}|) = (30, 40)$ としたときの、長さに関する分布の様子を図 2 に示す。 $(T, |\mathcal{A}|) = (30, 40)$ と

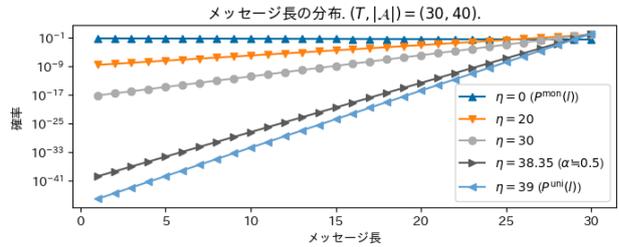


図 2 メッセージ長の分布 (($T, |\mathcal{A}|$) = (30, 40)).

いう超パラメータは、先行研究において典型的に選択されているものであり [2, 8, 9]、この分野の実態により近い。図 2 から分かるように、 $P^{\text{uni}}(l)$ は更に右側に偏っている。 $P^{\text{mon}}(l)$ は平坦な形状に見えるが、実際には依然として左側に偏っていて、右肩下がりの形状をしている。図 2 における $\eta = 38.35$ は、メッセージ長罰則項の重み $\alpha = 0.5$ に概ね対応している。これは [2] が選択した値と同じであり、この値で Zipf 短縮の創発に成功している。 $P^{\text{int}}(l; \eta = 38.35)$ が尚も右側に著しく偏っていることを踏まえると、

6) モンキータイピング系列が Zipf 則や Zipf 短縮に従うことは既に知られており [14, 5]、同様のことが起きている。

[2] が $\alpha = 0.5$ という重みで Zipf 短縮の創発に成功したという事実は驚くべきことである。これまで、創発言語の反 Zipf 短縮性はバイアスのせいであると考えられてきたが、 $P^{\text{uni}}(l)$ よりもわずかに偏りが緩和されただけの分布で Zipf 短縮を創発できるのなら、むしろエージェントは短いメッセージを選好する望ましいバイアスをもっているのではないかとさえ思える。ただし、あるメッセージ m_1 が $P^{\text{int}}(m; \eta)$ からサンプリングされる確率は、あるメッセージ m_2 が $P^{\text{int}}(m; \eta)$ からサンプリングされる確率の $(|\mathcal{A}| - \eta)^{|m_2| - |m_1|}$ 倍である。つまり、個々のメッセージ同士を比較する限りにおいては、 $P^{\text{int}}(m; \eta)$ は短いメッセージを選好する分布であるともいえる。

3.3 beta-VAE の目的関数との類似性

ゲームの目的関数に暗黙の事前分布が存在するという事は、ゲームの最適化は、ある種の生成モデルに関する変分推論と見做せるかもしれない。この発想に基づき、本節ではゲームの目的関数と beta-VAE [15] の目的関数との類似性について議論する。入力 i を観測変数 x 、メッセージ m を潜在変数 z 、送信者 $P_{\theta_s}(m | i)$ を符号化器 $q_\phi(z | x)$ 、受信者 $P_{\theta_r}(i | m)$ を復号化器 $p_\theta(x | z)$ 、事前分布 $P^{\text{int}}(m; \eta)$ を潜在変数の事前分布 $p_\theta(z)$ だと思ふことにすれば、シグナリングゲームの構造は VAE [16] のそれに類似している。そこで、シグナリングゲームの記法で以下のような beta-VAE の目的関数を考える：

$$\mathcal{J}_{\text{vae}}(\beta) := \mathbb{E}_{P_I(i)} [\mathbb{E}_{P_{\theta_s}(m|i)} [\log P_{\theta_r}(i | m)] - \beta \text{KL}(P_{\theta_s}(m | i) \| P^{\text{int}}(m; \eta))]. \quad (12)$$

beta-VAE では通常の VAE の目的関数を拡張し、KL ダイバージェンス項を β で重み付けする。 $\mathcal{J}_{\text{vae}}(\beta)$ は以下のように変形できる (証明は付録 A 参照)：

$$\mathcal{J}_{\text{vae}}(\beta) = \mathcal{J}_{\text{game}}^{\text{LP}}(\beta \log\{|\mathcal{A}| - \eta\}) + \beta \mathbb{E}_{P_I(i)} [\mathcal{H}(P_{\theta_s}(m|i))] + \text{const}. \quad (13)$$

式 13 の第 1 項は、メッセージ長罰則項の重みを $\alpha = \beta \log\{|\mathcal{A}| - \eta\}$ としたときのゲームの目的関数 $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ に一致している。第 2 項は送信者 P_{θ_s} の **エントロピー最大化項** になっている。このような、ゲームの報酬の期待値とエントロピー最大化項が合わさった目的関数を用いる設定は、しばしば **最大エントロピー強化学習** と呼ばれる。⁷⁾ 実は、創発言語

7) 変分推論と最大エントロピー強化学習が同一視できることは control-as-inference [17] の一例としても理解できる。

の研究でも **エントロピー正則化項** [18] という補助目的関数が頻繁に用いられる。エントロピー正則化項は、エントロピー最大化項と同様の効果を持つ。ただし、ややアドホックな手法であり、エントロピー最大化項とは一般には一致しない。以上のことから、ゲームはアドホックな補助関数を伴いつつ、暗に変分推論をしていると思われる。⁸⁾

3.4 以上の結果から得られる示唆

3 節では、まずゲームの目的関数 $\mathcal{J}_{\text{game}}, \mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ が特定の事前メッセージ分布 $P^{\text{int}}(m; \eta)$ を仮定していることを示した。続いて、事前メッセージ分布の形状に着目し、その形状から反 Zipf 短縮の一因を垣間見た。⁹⁾ 最後に VAE との類似性に触れ、ゲームが (アドホックな形で) 暗に生成モデルの変分推論をしていることが示唆された。本稿では、Zipf 短縮という特定の性質を取り上げて議論したが、ここでの考え方は、シグナリングゲームを用いる研究一般に適用可能なものである。従って、ゲームを用いる際には、暗に想定されている事前分布と、それがもたらし得る影響を意識することが重要になるだろう。

4 最後に

本稿の目的は、シグナリングゲームにおいて創発する言語が Zipf 短縮に従わないという知見に着目し、その原因の一端を示唆することであった。具体的には、目的関数が暗にもつ事前メッセージ分布に着目し、それが反 Zipf 短縮の一因になり得ることを示唆した。ゲームにおいて可変長メッセージを用いるのはごく自然に思える。しかし実は、先行研究の殆どは固定長メッセージを用いている。素朴に可変長にしても不自然にメッセージが長くなること、メッセージ長罰則項は学習を不安定にし得ること、罰則化項の (進化) 言語学的な意味が不明瞭なこと等が可変長が忌避される要因ではないかと推測している。ゲームの最適化に潜むアーティファクトを 1 つずつ洗い出して解決していくことが、今後の言語創発研究を豊かなものにしていくだろう。

8) また、beta-VAE と **情報ボトルネック法** (IB) [19] の目的関数にも密接な関係があることが知られている [20, 21]。IB を用いた言語創発研究としては [22, 23] などがある。ただし、これらは 1 メッセージ 1 記号の設定を用いているため、今回のような Zipf 短縮関連の議論を適用できない。

9) ただし、 $P^{\text{uni}}(m)$ が創発言語を“余計に長くする”ことが分かって、反 Zipf 短縮になる理由はそれほど自明でない。恐らく、NN のバイアスやエントロピー正則化のアドホックさなど、様々な要因が絡み合い反 Zipf 短縮が生じる。このことについて、より厳密に議論するのが今後の課題である。

謝辞

執筆に際してコメントを寄せてくださった宮尾祐介先生、李凌寒さんに感謝申し上げます。

参考文献

- [1] Angeliki Lazaridou and Marco Baroni. Emergent multi-agent communication in the deep learning era. **CoRR**, Vol. abs/2006.02419, , 2020.
- [2] Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. Anti-efficient encoding in emergent communication. In **Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada**, pp. 6290–6300, 2019.
- [3] George K. Zipf. **The psycho-biology of language**. Houghton Mifflin, 1935.
- [4] Jasmeen Kanwal, Kenny Smith, Jennifer Culbertson, and Simon Kirby. Zipf’s law of abbreviation and the principle of least effort: Language users optimise a miniature lexicon for efficient communication. **Cognition**, Vol. 165, pp. 45–52, 2017.
- [5] 田中久美子. 言語とフラクタル. 東京大学出版会, 2021.
- [6] David K. Lewis. **Convention: A Philosophical Study**. Wiley-Blackwell, 1969.
- [7] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. **Neural Comput.**, Vol. 9, No. 8, pp. 1735–1780, 1997.
- [8] Mathieu Rita, Rahma Chaabouni, and Emmanuel Dupoux. ”lazimpa”: Lazy and impatient neural agents learn to communicate efficiently. In **Proceedings of the 24th Conference on Computational Natural Language Learning, CoNLL 2020, Online, November 19-20, 2020**, pp. 335–343. Association for Computational Linguistics, 2020.
- [9] Ryo Ueda and Koki Washio. On the relationship between zipf’s law of abbreviation and interfering noise in emergent languages. In **Proceedings of the ACL-IJCNLP 2021 Student Research Workshop, ACL 2021, Online, JULI 5-10, 2021**, pp. 60–70. Association for Computational Linguistics, 2021.
- [10] George K. Zipf. **Human Behaviour and the Principle of Least Effort**. Addison-Wesley, 1949.
- [11] John Schulman, Nicolas Heess, Theophane Weber, and Pieter Abbeel. Gradient estimation using stochastic computation graphs. In **Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada**, pp. 3528–3536, 2015.
- [12] Ronald J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. **Mach. Learn.**, Vol. 8, pp. 229–256, 1992.
- [13] Bengt Sigurd, Mats Eeg-Olofsson, and Joost Van Weijer. Word length, sentence length and frequency – zipf revisited. **Studia Linguistica**, Vol. 58, No. 1, pp. 37–52, 2004.
- [14] George A. Miller. Some effects of intermittent silence. **The American Journal of Psychology**, Vol. 70, No. 2, pp. 311–314, 1957.
- [15] Irina Higgins, Loïc Matthey, Arka Pal, Christopher P. Burgess, Xavier Glorot, Matthew M. Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. In **5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings**. OpenReview.net, 2017.
- [16] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In **2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings**, 2014.
- [17] Sergey Levine. Reinforcement learning and control as probabilistic inference: Tutorial and review. **CoRR**, Vol. abs/1805.00909, , 2018.
- [18] Ronald J. Williams and Jing Peng. Function optimization using connectionist reinforcement learning algorithms. **Connection Science**, Vol. 3, No. 3, pp. 241–268, 1991.
- [19] Naftali Tishby, Fernando C. N. Pereira, and William Bialek. The information bottleneck method. **CoRR**, Vol. physics/0004057, , 2000.
- [20] Alessandro Achille and Stefano Soatto. Information dropout: Learning optimal representations through noisy computation. **IEEE Trans. Pattern Anal. Mach. Intell.**, Vol. 40, No. 12, pp. 2897–2905, 2018.
- [21] Alexander A. Alemi, Ian Fischer, Joshua V. Dillon, and Kevin Murphy. Deep variational information bottleneck. In **5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings**. OpenReview.net, 2017.
- [22] Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. Communicating artificial neural networks develop efficient color-naming systems. **Proceedings of the National Academy of Sciences**, Vol. 118, No. 12, p. e2016569118, 2021.
- [23] Mycal Tucker, Roger P. Levy, Julie Shah, and Noga Zaslavsky. Trading off utility, informativeness, and complexity in emergent communication. In **Advances in Neural Information Processing Systems, 2022**.

A 証明

本節では、スペースの都合で本文には掲載できなかった証明を示す。

式 6 ・ 式 8 ・ 式 10 の証明. 本文で述べたように、 $P^{\text{int}}(m; \eta)$ は $P^{\text{uni}}(m)$ 及び $P^{\text{mon}}(m)$ を包含している。また、 $\alpha = 0$ のとき $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha) = \mathcal{J}_{\text{game}}$ であるから、 $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ は $\mathcal{J}_{\text{game}}$ を含む。従って、式 10 のみを示せばよい。

$$\begin{aligned}
& \nabla \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}^{\text{int}}(i, m; \eta)] \\
&= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\nabla \log P_{\theta_r}^{\text{int}}(i, m; \eta) + \{\log P_{\theta_r}^{\text{int}}(i, m; \eta) - b(i)\} \nabla \log P_{\theta_s}(m|i)] \\
&= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\nabla \log P_{\theta_r}(i|m) + \nabla \log P^{\text{int}}(m; \eta) + \{\log P_{\theta_r}(i|m) + \log P^{\text{int}}(m; \eta) - b(i)\} \nabla \log P_{\theta_s}(m|i)] \\
&= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\nabla \log P_{\theta_r}(i|m) + \{\log P_{\theta_r}(i|m) + \log P^{\text{int}}(m; \eta) - b(i)\} \nabla \log P_{\theta_s}(m|i)] \\
&= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\nabla \log P_{\theta_r}(i|m) + \{\log P_{\theta_r}(i|m) - |m| \log(|\mathcal{A}| - \eta) - \log Z' - b(i)\} \nabla \log P_{\theta_s}(m|i)] \\
&= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\nabla \log P_{\theta_r}(i|m) + \{\log P_{\theta_r}(i|m) - |m| \log(|\mathcal{A}| - \eta) - b'(i)\} \nabla \log P_{\theta_s}(m|i)] \\
&= \nabla \mathcal{J}_{\text{game}}^{\text{LP}}(\log\{|\mathcal{A}| - \eta\}).
\end{aligned}$$

1 つ目の等式は式 2 と同様の式変形によって得られ、2 つ目の等式は $P_{\theta_r}^{\text{int}}(i, m; \eta)$ の定義から従い、3 つ目の等式は $\nabla P^{\text{int}}(m; \eta) = \mathbf{0}$ であることから従い、4 つ目の等式は $P^{\text{int}}(m; \eta)$ の定義から従い、5 つ目の等式は新たなベースライン関数 $b'(i) = b(i) + \log Z'$ を置くことで得られ、最後の等式は式 4 と同様の式変形によって得られる。

式 13 の証明. 式 13 は以下のように証明できる。

$$\begin{aligned}
\mathcal{J}_{\text{vac}}(\beta) &= \mathbb{E}_{P_I(i)} [\mathbb{E}_{P_{\theta_s}(m|i)} [\log P_{\theta_r}(i|m)] - \beta \text{KL}(P_{\theta_s}(m|i) \| P^{\text{int}}(m; \eta))] \\
&= \mathbb{E}_{P_I(i)} [\mathbb{E}_{P_{\theta_s}(m|i)} [\log P_{\theta_r}(i|m)] - \beta \mathbb{E}_{P_{\theta_s}(m|i)} [\log P_{\theta_s}(m|i) - \log P^{\text{int}}(m; \eta)]] \\
&= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}(i|m) - \beta \log P_{\theta_s}(m|i) + \beta \log P^{\text{int}}(m; \eta)] \\
&= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}(i|m) + \beta \log P^{\text{int}}(m; \eta)] - \beta \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_s}(m|i)] \\
&= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}(i|m) + \beta \log P^{\text{int}}(m; \eta)] + \beta \mathbb{E}_{P_I(i)} [\mathcal{H}(P_{\theta_s}(m|i))] \\
&= \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}(i|m) - \beta |m| \log(|\mathcal{A}| - \eta) - \log Z'] + \beta \mathbb{E}_{P_I(i)} [\mathcal{H}(P_{\theta_s}(m|i))] \\
&= \mathcal{J}_{\text{game}}^{\text{LP}}(\beta \log\{|\mathcal{A}| - \eta\}) + \beta \mathbb{E}_{P_I(i)} [\mathcal{H}(P_{\theta_s}(m|i))] + \text{const.}
\end{aligned}$$

1 つ目の等式は $\mathcal{J}_{\text{vac}}(\beta)$ の定義 (式 12) そのものであり、2 つ目の等式は KL ダイバージェンスの定義から従い、3 つ目・4 つ目の等式は期待値の線形性から従い、5 つ目の等式はエントロピーの定義から従い、6 つ目の等式は $P^{\text{int}}(m; \eta)$ の定義から従い、7 つ目の等式は $\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ の定義及び $\nabla \mathbb{E}[\log Z'] = \nabla \log Z' = \mathbf{0}$ から従う。

B メッセージ長分布をガンマ分布としてゲームの目的関数を逆算した場合

自然言語における単語長分布はガンマ分布に従うことが知られている [13]。では、メッセージ長をガンマ分布として、そこからゲームの目的関数を逆算したらどのような形になるのだろうか。まず、メッセージ長がガンマ分布に従うと仮定する: $P^{\text{gam}}(l; k, \lambda) := \frac{1}{Z'''} l^{k-1} \exp(-\lambda l)$ 。 $k > 0, \lambda > 0$ は超パラメータであり、 Z''' は適当な正規化項である。このとき、 $P^{\text{gam}}(l; k, \lambda)$ に無矛盾なメッセージ分布として、以下が挙げられる:

$$P^{\text{gam}}(m; k, \lambda) := \frac{1}{Z'''} |m|^{k-1} \exp(-\{\lambda - \log(|\mathcal{A}| - 1)\} |m|).$$

ただし、 $Z''' = \frac{Z''}{\log(|\mathcal{A}| - 1)}$ である。 $P^{\text{gam}}(m; k, \lambda)$ と受信者 $P_{\theta_r}(i|m)$ を合わせて、 i, m の同時分布 $P_{\theta_r}^{\text{gam}}(i, m) := P_{\theta_r}(i|m) P^{\text{gam}}(m; k, \lambda)$ を定義し、新たなゲームの目的関数 $\mathcal{J}_{\text{game}}^{\text{gam}}(k, \lambda)$ を考える:

$$\mathcal{J}_{\text{game}}^{\text{gam}}(k, \lambda) := \mathbb{E}_{P_I(i), P_{\theta_s}(m|i)} [\log P_{\theta_r}^{\text{gam}}(i, m)].$$

ここで、 $\mathcal{J}_{\text{game}}^{\text{gam}}(k, \lambda)$ の微分は以下の形で得られる:

$$\nabla \mathcal{J}_{\text{game}}^{\text{gam}}(k, \lambda) = \mathbb{E} [\nabla \log P_{\theta_r}(i|m) + \{\log P_{\theta_r}(i|m) + (k-1) \log(|m|) - (\lambda + \log(|\mathcal{A}| - 1)) |m|\} \nabla \log P_{\theta_s}(m|i)].$$

$\mathcal{J}_{\text{game}}^{\text{LP}}(\alpha)$ 同様の長さ に比例する罰則に加え、長さの対数に比例する報酬が追加されたような形になる。