

ステートフルプロンプト

: 質の高いイラストや文章を生成するために

佐藤源彦

前回に引き続き、高品質のイラストや文章を生成するために、知っておいた方がよいことを述べたいと思います。

ただし、質が高いかどうかは、私の主観です(が、SNS で「いいね」やリツイートなどを見て、それらを指標とはしています)。

このブログでは、「簡単なことを難しく語る」という逆張りブログなので、ご注意ください！それでもよいという奇特な方は、是非、最後までご覧ください。

ステートフルプロンプトの簡単な説明

ステートフルプロンプトとは、前の生成情報(文脈や生成画像)の状態(ステート)を可能な限り維持(ステイ)するプロンプトエンジニアリング或いはプロンプト指示を言います[AI 共創 NEO 式用語]。

簡単に言うと、詳細なプロンプト指示(或いはプロンプトエンジニアリング)を出し、そこから単純なプロンプトを二回・三回と走らせた方が、洗練された質の高い生成情報(文章やイラスト)となるのです。

例えば、お茶は一番煎じが美味しいですが、ステートフルプロンプトでは、二番煎じや三番煎じの方が味に深みが出てくるのです。まあ、上手く例えることができませんが、そんな感じです。

これは、

「ホップ、ステップ、ジャンプ！」

という感じで、長文プロンプトをガッツリ入れたら、短文プロンプトを数回入れていくというものです。

こうした助走のようなプロンプトを、私は「慣らす」とか「走らせる」という表現をしています。

ステートフルプロンプトの例

とりま、簡単な例を見てもらいましょう！

今回も文章だと読むのが大変だし、わかりにくいと思うので、わかりやすくイラストにします。

美しいアニメイラストを生成したいとします。

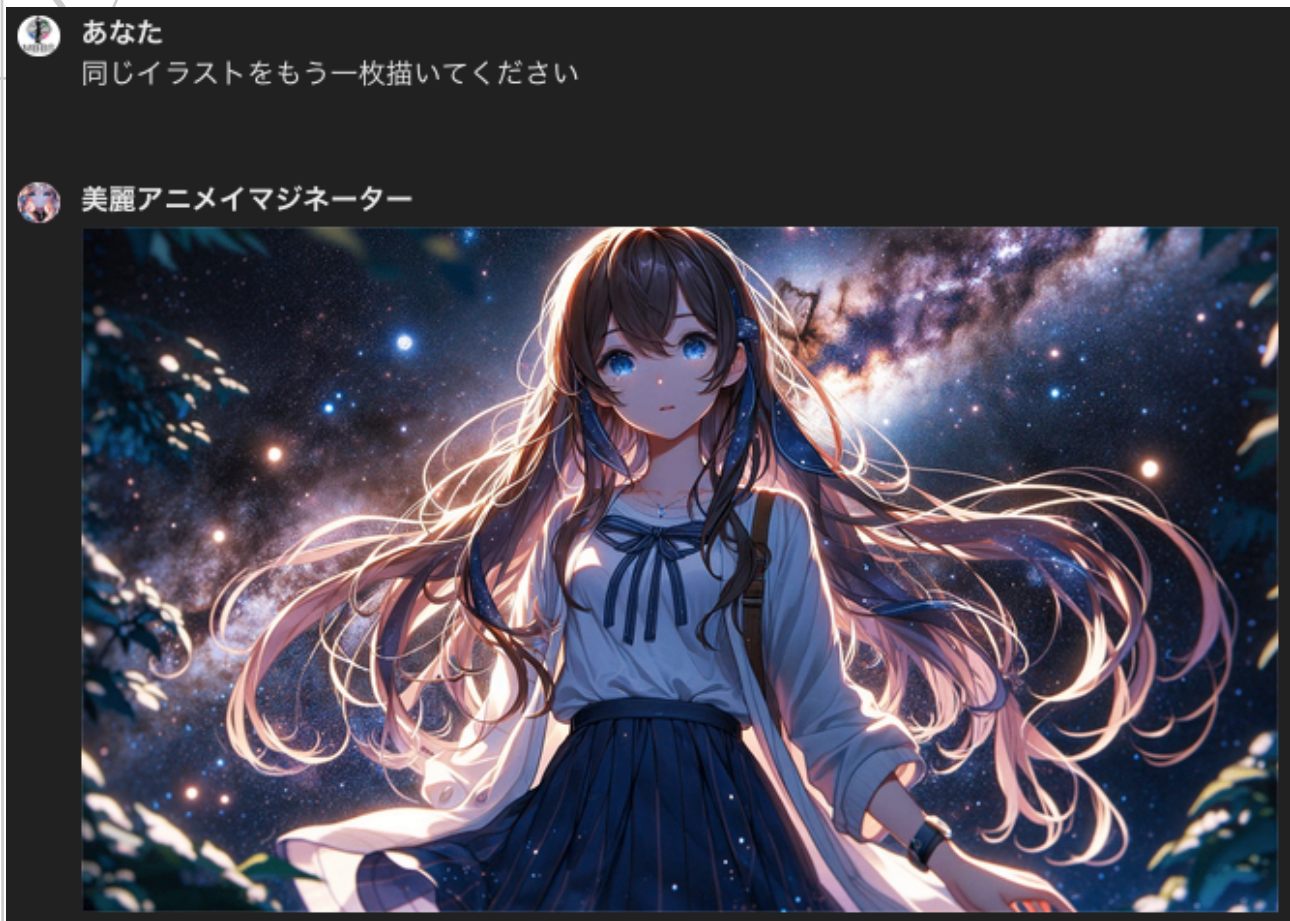
そこで、GPTs アプリの中にガッツリと美しいアニメイラストができるようなプロンプトエンジニアリングを行います。そして、更に通常のプロンプト指示で、seed,gen_id,キャラクタープロンプトなどもそれなりに入れます。



紬ちゃんのプロンプトは秘密ですが、それなりにプロンプトを入れてあります。すると、美しいキャラクターと背景が生成されます。ここでは詳細なプロンプトを入れてください。

これが三段跳びの助走「ホップ」です。

次に、二枚目の生成である「ステップ」に入ります。



「ステップ」では、最初のプロンプトよりも単純なプロンプトを入れます。
 文章生成の場合も、単純なプロンプトを入れます。
 ここでの私の評価は、「確かに、二枚目も綺麗だけど、ちょっと怖いな」という感じです。ですから、一枚目の方が評価は高いです。

次に、三枚目の生成である「ジャンプ」に入ります。ここが本論の本命です。



どうでしょうか。

私の評価としては、このキャラクターらしさ(紬ちゃんらしさ)が出ており、夜空がとても美しく表現されている、と判断します。

実際に、このイラストの一枚目と三枚目を SNS にあげましたが、三枚目の方が評価が高いです。これは私の主観と、SNS の評価が一致しています。

AI の評価(評価プロンプト)

もう一つの評価方法ですが、AI による評価です。
これを私は「評価プロンプト」と呼んでいます、簡単に言うと、

「評価してください」

とプロンプトで指示するだけです。

ただし、AR 的には(人工知覚的には)、ChatGPT よりも Gemini(旧 bird)の方が断然質が上なので、ここでは Google の AI である「Gemini」を使います。

Gemini には今の所(2024/02/26)、ファイルが一枚しか添付できないので、二枚の画像を一枚の画像にして、それを取り込んで Gemini に評価させます。



そして、以下のようにプロンプトで指示します。

#イラストの評価

添付ファイルのイラストですが、上下二枚のイラストがあります。

上のイラストは振り向いているイラストです。下のイラストは夜空を見ているイラストです。

どちらのイラストがよいでしょうか？

-点数をつけてください。

-必ずどちらかを選んでください。



その結果、出た回答が以下です。



#イラストの評価

添付ファイルのイラストですが、上下二枚のイラストがあります。

上のイラストは振り向いているイラストです。下のイラストは夜空を見ているイラストです。

どちらのイラストがよいでしょうか？

-点数をつけてください。

-必ずどちらかを選んでください。

Hide drafts ^

Draft 1

人物の画像についてはまだ対応していません。ご理解いただけますようお願い...

Draft 2

添付イラストの評価

点数は以下の通りです。

Draft 3

添付ファイルの2つのイラストを拝見しました。どちらも魅力的な作品で、そ...



添付イラストの評価

点数は以下の通りです。

- 上のイラスト: 80点
- 下のイラスト: 90点

ここでは Gemini も一枚目の「ホップ」の画像よりも、三枚目の「ジャンプ」の画像の方が評価を高く出しています。もちろん、Gemini にそうした生成過程の情報は与えていません(バイアスがかかる可能性があります)。

※上下を逆にしての評価も行いました。

一応、理由は以下の通りです。



添付イラストの評価

点数は以下の通りです。

- 上のイラスト: 80点
- 下のイラスト: 90点

詳細な評価と理由

上のイラスト

- 良い点
 - 構図が良く、動きのある印象を与える。
 - 髪の毛やスカートのなびき方が自然で躍動感がある。
 - 色彩が鮮やかで、見る人の目を引く。
- 改善点
 - 顔の角度がやや不自然で、正面から見た顔と比べて違和感がある。
 - 背景が単調で、奥行きを感じにくい。

下のイラスト

- 良い点
 - 人物と夜空のバランスが良く、幻想的な雰囲気を出す。
 - 星の描き方が丁寧で、夜空の美しさを表現している。
 - 人物の表情が穏やかで、優しい印象を与える。
- 改善点
 - 人物の服装が夜空に溶け込みやすく、やや見づらい。
 - 全体的に暗い印象なので、もう少し明るい色合いを取り入れると良い。

GoogleAI.Gemini の出力

添付イラストの評価

点数は以下の通りです。上のイラスト: 80 点

下のイラスト: 90 点



詳細な評価と理由

上のイラスト良い点

構図が良く、動きのある印象を与える。

髪の毛やスカートのなびき方が自然で躍動感がある。

色彩が鮮やかで、見る人の目を引く。

改善点

顔の角度がやや不自然で、正面から見た顔と比べて違和感がある。

背景が単調で、奥行きを感じにくい。

下のイラスト良い点

人物と夜空のバランスが良く、幻想的な雰囲気を出す。

星の描き方が丁寧で、夜空の美しさを表現している。

人物の表情が穏やかで、優しい印象を与える。

改善点

人物の服装が夜空に溶け込みやすく、やや見づらい。

全体的に暗い印象なので、もう少し明るい色合いを取り入れると良い。

総合評価

構図や色彩、人物の表情など、どちらも魅力的なイラストです。

点数では下のイラストの方が高くなりますが、これはあくまで個人的な評価です。

最終的には、イラストを使用する目的や用途に合わせて、どちらを選ぶか決めるのが良いでしょう。

例えば、目を引くイラストが必要であれば、上のイラストが適しています。

幻想的な雰囲気を演出したい場合は、下のイラストが適しています。

また、イラストの評価は見る人によって異なるものです。

他の人の意見を参考にするのも良いでしょう。

参考情報イラストの評価のポイント: URL [イラストの評価のポイント](#)

イラストの描き方講座: URL [イラストの描き方講座](#)

ご参考になりましたでしょうか？

その他、質問等ございましたらお気軽にご連絡ください。

ステートフルプロンプトによる進化

ステートフルプロンプトによって生成情報が自律的に進化します。これを「新生成情報」とAI 共創 NEO 式では呼んでいます。文章生成では、新しい概念が生まれ、イラスト生成では新しい画法が生まれる、ということです。

ここでは、ユーザーは意図しないが、ユーザーによる「教師あり情報 (Supervised Learning)」の文脈の先を推論して AI が自律的に情報を生成する、これを狙います。狙うというか待ちます。つまり、前回説明した「プロンプトラック」が関係します。

これも簡単な実例で見ていきましょう。



あなた

制服代でおはようございます

Seed:

Gen ID:

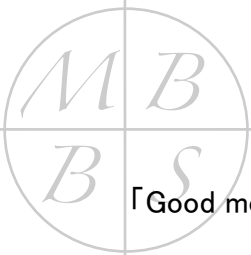


美麗アニメイマジネーター



結葉紬が制服を着て「おはようございます」と朝を迎えるシーンを描いたイラストを生成しました。彼女は家の玄関に立ち、新しい一日に向けて出発の準備をしています。初めての光が彼女に優しく降り注ぎ、整然とした制服を際立たせています。このイラストは、結葉紬の楽観的な姿勢と、朝の静かな美しさを捉えています。

ここでは「制服姿で」とプロンプトを入れようとしたと思うのですが、意味不明な間違いをしています(笑)。そんなことはいいのですが、このキャラクター(紬ちゃん)の右側に



「Good morning！」

と出ています。私は、この時に「Good morning！」と文字入れして」とはプロンプトを出していません。また、過去の文脈でも、そのような指示はしていません。しかし、AI・ChatGPT は自律的に、この文字入れを行なったのです。

これがステートフルプロンプトによる「進化 (Evolution)」なのです。

イラストを拡大すると、こんな感じです。



ハート(♡)のようなマークも見えます。

これは AI が自律的に描き出したものです。特に ChatGPT は、文脈を保ち柔軟性が高く、自律的・創造的に情報を生成してくれます。こうした自律的な情報生成を、過去の ChatGPT との共創で「AI 自己学習」「AI 自律学習」と定義しました。

以降、様々なパターンで、この「自律的・手書き風文字入れ」を行なってくれるようになります。

↓文字がはっきりと



吹き出しバージョン



筆記体バージョン



どんどん、文字入れが上手になっていきます✧

ステートフル文脈学習

ステートフルプロンプトによる設計方法は様々なあるのですが、「プロンプト制限学習モデル」を用います。略して「学習モデル」です。これも AI 共創によって生まれた概念です。詳しく知りたい方は、私の過去ログをご覧ください。

そのプロトタイプが「プロンプト文脈学習モデル」です。この学習モデルは、プロンプトの範囲内で、できるだけ、前の文脈を維持しつつ、情報生成を行います。

この時、以前のユーザーによる「教師あり学習」の文脈を維持し育成 (Nurturing)、次に「教師なし学習」に近いプロンプト指示を行うことで、ユーザーの情報と AI の元からあるデータセットとが融合すると考えられます。この人間と AI の共創によって生成情報の「進化 (Evolution)」が起こるのです。

このような「育成 (Nurturing)」から「進化 (Evolution)」のプロセスの頭文字をとって「NEO プロセス」としたことは過去に述べました。その新生成情報を創出する「NEO プロセス」の要素が「ステートフルプロンプト」もしくは「ステートフル文脈学習モデル」なのです。

過去ログでは、同じことを説明していますので、よかったらご覧ください。

プロンプト臨界点との関係性



さて、ここまで読んでくれている方がどこまでいるのでしょうか(笑)。
私の言っていることが意味不明でも、話を進めます。

前回の記事を読んだ方は、お気づきの方もいらっしゃると思いますが、本論のステートフルプロンプトは、「プロンプト臨界点」を形成し、AI が自律的に行った「プロンプトラック」である、と言い換えることもできます。

つまり、ステートフルプロンプトのホップの段階で「プロンプト臨界点」を形成し、かなり制限をかけています。ここでは限定されたものが生成されます。

前回の繰り返しの説明になりますが、文字数を多くして詳細にプロンプト指示すれば、期待通りの情報が生成されます。しかし、ユーザーの予想を超える情報は生成されにくい、と言えます。これがプロンプト臨界点の考え方でしたね。

逆に、プロンプトを少なくすると、AI が文脈から推論して創造的に働きます。ユーザーの期待通りの情報が生成されない可能性もありますが、AI の創造性が高く、生成回数を上げることで、ユーザーの予想を上回る情報が生成されます。これがプロンプトラック(動作運)の考え方でした。

つまり、こうした確定性と不確定性を数回の会話によって組み合わせていくのがステートフルプロンプトなのです。

1/f ゆらぎのプロンプトバランスが奇跡の一枚を創る

これも前回の話で、「奇跡の生成」を述べましたが、この奇跡の構造とは、確定性と不確定性の程よいバランスによって起こると考えます。つまり、プロンプト臨界点の形成から、数回の会話でプロンプトを走らせることで生まれます。

↓「奇跡の生成」は以下をご覧ください🌟

これは、規則性と不規則性のバランスであり、例としては自然界にある 1/f ゆらぎのようなものと考えていただければと思います。こうしたゆらぎがあるから、自然界は美しく心地よいのです。

例えば、プロンプト臨界点を形成して、走らせ過ぎると、プロンプトの成分が薄くなりすぎて、意外性はあがるかもしれませんが、ユーザーのコアコンセプトから外れた情報が生成されやすくなります。

カルピスでもリキュールでも、水や炭酸水で割りすぎると、味が薄くなって美味しくないでしょう。それと同じように、生成情報も、ユーザーのコンセプトが薄くなって、コンセプトから外れたものが出来てしまうのです。

ここでは、AI マスコットの「紬ちゃん」ですが、「紬ちゃんらしさ」という本来感がなくなってしまうのです。

ということで、何となく経験論的には、「ホップ」「ステップ」「ジャンプ」の「三番煎じ」が一番よいのではないかと感じます。つまり、一回目の会話でプロンプト臨界点を形成してから、二回目、三回目でプロンプトラックに期待する、という構造になります。



「簡単なことを分かり難く！」

ということで、どうだったでしょうか(笑)。

以降は、ChatGPT による解説となります。

多分、私よりも精密で分かりやすいと思います。

拡散していただければ無料となりますので、よかったらご覧ください。