

### 1. 担当 PM

田中 邦裕（さくらインターネット株式会社 代表取締役社長）

### 2. クリエータ氏名

三林 亮太（兵庫県立大学大学院情報科学研究科 博士後期課程）

### 3. 委託金支払額

2,736,000 円

### 4. テーマ名

ラップバトル対話システムの開発

### 5. 関連 Web サイト

なし

### 6. テーマ概要

本プロジェクトでは、コンピュータの処理能力と豊富な語彙を用いて、ラップバトル対話システムを開発する。ラップバトルの対話を行うには、相手のバースに対して返答となるバースを生成する必要がある。返答バース生成はテキストベースと音声ベースの生成を行う。本システムの軸となるのはテキストベースでのバース生成である。音声ベースの生成はこのテキストを音声に変換することで生成を行う。

本プロジェクトでは、バース生成のみならず、バース生成システムを用いたこれまでになかった新しいラップバトルへの開拓を目的としている。本対話システムは実際にラップバトルの大会にエントリーし、一人のラッパーとして出場させることを目指した。また、本システムを用いることにより、これまでになかったラップバトルコンテンツの作成が期待される。加えて、本来人同士で行われるラップバトルをシステム対システム、人对システム、という新しいラップバトルの形にて実現可能である。

## 7. 採択理由

本プロジェクトでは、ラップバトルが本当に好きな提案者が、自らの計算機科学の能力を活かして、計算機を1人のラッパーとしてラップバトルに出場させるという野心的なもので、元来人間がやるべきものと思われていたエンターテインメント分野における、新星が現れたものと考えます。

ラップバトルとは極めてクリエイティビティの高い競技であり、ライム（韻）の良さと、アンサー（返答）のバリエーション、そしてフロー（リズム）の印象という3つの要素を、相手の即興ラップ（バース）を反映しながら考えるというのは、プロでも難しい。

他方、ライムの分解については母音の判定であり、バースの選択はボキャブラリーの探索であり、実は計算機が得意とするところであるのが重要なポイントである。

プロジェクトにおいては、人間の声を解析し、相手のバースを構造的に理解して、文脈的にも一致するアンサーを返す一連の動作を実装する必要があり、既存のAPIやライブラリを組み合わせるだけではない相当な困難性も伺える。

このように、本人のこれを実現したいという高いモチベーション、計算機を利用することの合理性、そして乗り越えないといけな未踏な困難性、何よりこのプロジェクトの成果発表会で実施されるであろうラップバトルを期待して採択することとした。

## 8. 開発目標

本プロジェクトの目標達成のために、開発期間ですべきことは、実戦デモシステム、バース生成システム、ラップ音声合成システムの作成の、大きく分けて3つである。

本プロジェクトの核となるシステムはバース生成システムである。これは入力バースに対して、返答バースを生成するテキストベースのシステムである。本プロジェクトでは、このバース生成システムを作成することを第一目標とした。次に、バーステキストをラップ音声に変換するシステムを作成することを目標とした。これは、生成された返答バースをラップ音声に変換するシステムである。最後に、本技術を利用するための実戦デモシステムを作成することを目標とした。これは、バース生成システムとラップ音声合成システムを用いた、実空間でラップバトルが体験できるデモシステムである。

## 9. 進捗概要

本プロジェクトでは、まずラップバトルにおいて必要とされる「ライム」と「アンサー」を考慮したバース生成システムを開発した。具体的には、深層学習モデルを用いてラップバトルコーパスからラップ生成を行う手法を提案した。

ラップバトルコーパスは、ラップバトル動画内のバースを文字起こしして構

築した。このため、クラウドソーシングを活用し、ラップバトル動画の文字起こしを行った。ラップバトルコーパスは約 1,780 件であり、バース単位で計算すると 9,749 バースを含んでいる (図 1)。

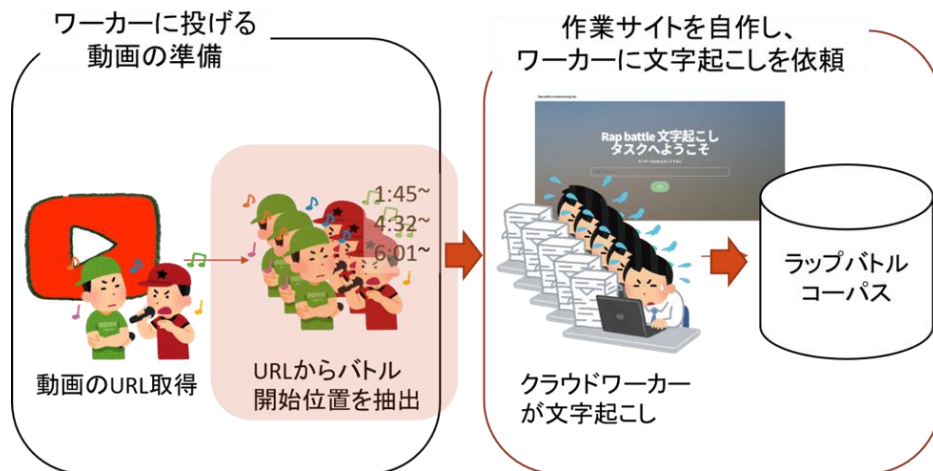


図 1: クラウドソーシングの流れ

ライムを考慮したラップ生成手法では、逆向き生成による方法を採用した。Transformer Encoder-Decoder を用いて、逆方向からの文生成を行うことで、文末にライムを考慮したラップを生成する。また、アンサーを考慮したバース生成システムについても、アンサーの対応関係を表した学習データを作成し、独自の評価関数によるフィルターを加えて開発した (図 2)。これにより、相手の発言をふまえた返答をすることができるようになり、ラップバトルの楽しさをより一層高めることができると思われる。

## Encoder-Decoder

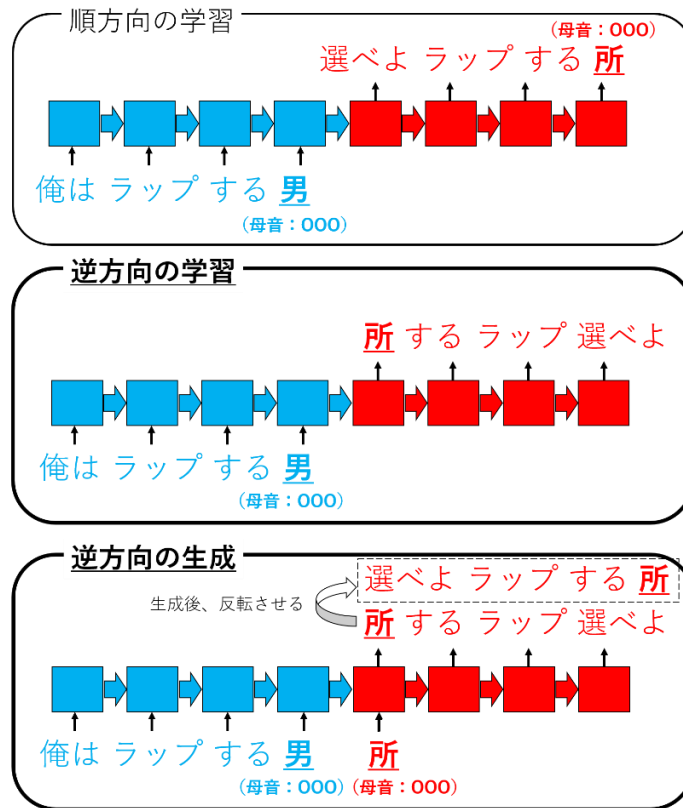


図 2 : Encoder-Decoder による文生成と逆向き生成

次にラップ音声合成システムであるが、Tacotron2 と Griffin-Lim アルゴリズムを用いて実装し、ラップ調の音声合成が可能となった。また、ラップバトルに合わせた高速な音声合成や、生成音声の強制的な推論切り上げなど、即興に対応した実装も行った。

成果報告会では、対応が間に合わず、CoeFont のラップ音声合成 API を一部利用したが、その後ラップ音声コーパスを自作し、完全に自作のラップ調音声合成を開発した。このラップ音声コーパスは、324 文のテキストに対してラップ調に読んだ音声データの集合である。ラップに特化した音声コーパスは存在しないため、自作する必要があった。本コーパスは、クリエイター自身がラップ調に読んだオリジナルのコーパスである。

コーパス用に読むテキストは、ITA コーパスと呼ばれる、合計 424 文からなる、音素バランスを考慮したパブリックドメインの日本語テキストデータから 324 文を選択した。今回、音声収録は専門家のアドバイスのもと、コンデンサマイクを用いて静かな部屋で収録を行った。収録を支援するツールとして、簡易の Web サイトを作成した。本サイトでは、324 文に対して読むべきテキストとその読み仮名を表示してくれる。また、音声の見本を聞くこともできる。録音には Audacity という録音のフリーソフトを用いて行い、サンプリング周波数は 44,100、チャンネルはステレオ、量子化は 32bit 浮動小数点に設定した。結果

として、324 文をラップ調にて読み上げたデータをもとに、文ごとの分割処理を行い、324 件のラップ音声コーパスを取得した。収録した音声データは Tacotron2 の学習で用いるために、前処理し、メルスペクトログラムに変換した。

最後に、人間とラップバトルが対戦できるインタラクティブなシステムである実戦デモシステムを開発した。このシステムは、ハードウェアとソフトウェアの両方で構成され、ハードウェアには PC、ミキサー、スピーカー、有線マイクを搭載した折りたたみ式台車が使用され、ソフトウェアにはラップバトルが可能な Web アプリケーションが使用されている。

ハードウェアは、持ち運びや人間とのインタラクティブ性を考慮した設計となっており、台車には予備バッテリーを搭載し、周囲に荷物が残らないようにカバーを設置した。また、ホイールをエラストマー素材に差し替えることで騒音を軽減し、新幹線や電車等の公共交通機関を利用する場合の持ち運びにも配慮した (図 3)。



図 3 : ハードウェアと構成要素

ソフトウェアでは Web アプリケーションを開発し、ユーザがマイクを持ち、実戦デモシステムと対峙する形式でのラップバトルを実現している。このシステムは、ユーザが入力したラップを取り込み、それに対する相手のラップをスピーカーから再生することで、リアルタイムでのラップバトルを可能にしている。

実戦デモシステムは、ラップバトルだけでなく、ヒップホップ文化の普及にも貢献することが期待されており、特に若者を中心に人気の高いヒップホップカルチャーを身近に感じられるようにすることが狙いとされている。今後は、実際のラップバトルイベントでの使用や、さらなる改良によるシステムの進化が期待される。

以上の開発により、世界で初めてラップバトルをコンピュータと対戦するこ

とが可能となった。この成果の特徴としては、これまで対人であったラップバトルが一人で完結することにある。例えば、これまで敷居が高かったラップバトルの練習が一人でおこなえるようになるなど、応用例はいくつか考えられる。加えて、AI 対 AI 同士のラップバトルも可能となり、鑑賞する競技でもあるラップバトルを実質無限に見ることができる。つまり本開発成果は、ラップバトルをラッパーなしで再現可能となったことを意味する。

## 10. プロジェクト評価

本プロジェクトでは、音声合成の精度や速度、音声変換の正確性など、多くのチャレンジがあったが、その結果、高品質で自然なラップ調の音声合成を実現し、即興に合わせた高速な音声合成も可能となった。さらに、ビートに合わせた音声変換も成功し、ラップバトルにおいて実用化が可能となった。このプロジェクトの成果は、ラップバトルに限らず、音声合成技術や自然言語処理技術の発展に大きな貢献をするものであり、その評価は非常に高いものと言える。また、自作したラップ音声コーパスにより、ラップに特化した音声コーパスの不足を補うこともできたため、今後の音声合成技術の発展にも貢献することが期待される。

さらに、HCI（ヒューマンコンピュータインタラクション）という観点において、人間のクリエイティビティをコンピュータによって拡張するという観点からも斬新である。単純な効率性ではなく、韻を踏むことの面白さであったり、型の制約による難易度であったり、人間の生活を豊かにするような分野にコンピュータを応用するという観点でも非常に面白い。また、ラップに合わせた音声合成も、素晴らしい結果であった。

なお 11 月の八合目会議においては、ラップ的に意味があり面白いレスポンスを得ることができなかつたり、音声認識がうまく行かず手入力で行わざるを得なかつたり、プロジェクトの方向性をどうするのか、プロジェクトマネージャーとしても悩む時期もあったが、年始にかけて上記のようなチューニングを行うことで、徐々にコンピュータがラップに応答する姿を見て、プロジェクト自体の成長を感じることができたことは思い出深い。

今後、プロジェクト期間中にお世話になった産業総合研究所の皆様の支援も得ながら、本人がラップバトルを博士課程進学とともにブラッシュアップすると聞いており、更なる発展が期待できるプロジェクトであった。

## 11. 今後の課題

今後の課題としては、本システムを API 化し、サービスとして運用することで、他のユーザがラップバトルシステムの構築を容易にすることが挙げられる。実際、システムの API 化はできているようであるが、GPU マシンが必要という制約があるため、運用コストがネックである。現在は GPU マシンを使わない手法で、いくつかの運用方法を検討しているようだ。また、ラップバトルにおける

言葉選びやフローの正確性など、人間の持つ表現力やセンスをAIが再現するためには、さらなるデータの収集やアルゴリズムの改良が必要である。そのため、今後もコーパスの充実やアルゴリズムの改善を行っていく必要がある。

展望としては、将棋AIや囲碁AIがその競技に対して、大きなインパクトを与えたように、ラップバトルにおいても大きなインパクトを与えることが期待される。将棋ではプロ棋士がAIを用いて対局の訓練をするなど、競技に対するひとつのパーツとして機能している。ラップバトルでもプロのラッパーがAIと対戦をしてライムの踏み方やアンサーの返し方などを訓練するような環境にまで持っていければ広がりは大いだと考える。

また、AI同士のラップバトルという新たなコンテンツにも注目しており、これまでは人間対人間であったのが、本プロジェクトによってAI対人間の対戦が可能にかつ、そこから更にAI対AIの対戦という応用が可能になった。ラップバトルは見る競技の側面が大きいため、AI対AIの対戦を見るということに大きな価値があると考えます。また、プログラミングのコンテストがあるように、各ユーザが作成したAIラッパー同士を戦わせることで、AIだけのラップバトル大会や、AIサイファーなどが実現可能ではないかと考える。