

07

いろいろな観点で欲しいものを見つけます

～グラフ索引型探索技術による絵本検索システム「ぴたりえ」～

どんな研究

「間違い混じりのあらすじを入力して絵本を探す」「好みの絵本と似た画風の作家による絵本を探す」など、**いろいろな観点で絵本を検索できるシステム**について紹介します。独自の検索技術に加え、ひらがな文の自然言語処理技術、子どもの語彙獲得についての知見を取り入れた学際的研究です。

どこが凄い

提案手法は、検索対象となる大量のデータを事前に整理して、似たもの同士が接続されたグラフ(ネットワーク)を構成しておくことで高速な検索を実現します。従来手法と異なり、メディア(文章、画像、音声など)の特性に依存しないため、**幅広いメディア**から抽出する特徴に対応して**高速化**を実現します。

めざす未来

病院での発音訓練に適した絵本探し、小学校での英語教育に適した英語絵本探しなど、**絵本を活用した様々なサービス実現**をめざすとともに、一般書籍や技術文書など**絵本以外のメディア**についても**多様な観点からの検索を実現**するための基盤となるような研究を行います。

絵本検索システム「ぴたりえ」



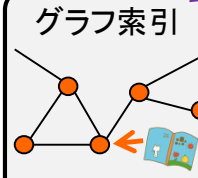
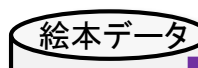
いろいろな観点で絵本を探せます。

観点1: あらすじ

観点2: お気に入りの絵本と内容が似ている

観点3: 画風が似ている (下記例)

検索に先立ちグラフ索引を用意(オフライン処理)



グラフ索引構築アルゴリズム

似たもの同士が接続されたグラフを構成し、索引として活用。

- 「スモールワールド性」:
任意の2頂点が少数ホップ以内に存在し、検索に適している。
- メディアの特性に依存しない工夫なので、文章画像音声など多様な対象に適用可能。

グラフ索引を用いて高速な検索を実現

検索例

入力した絵本と同作者の絵本はもちろん、似た画風の作家の絵本も見つかります。



クエリ



出力
クエリ絵本と異なる作者の絵本

引用させていただいた絵本:

しろくまちゃんのほっとけーき、わかやまけん作、こぐま社、1972年
うさこちゃんとうみ、ディックブルーナ作、福音館書店、1964年
(「雪の女王」は本スライドのため内製したものです)

様々な応用

応用例

構音(発音)障害の訓練に適した絵本の検索システム「ぴたりえメディカル」



サ行の発音をタ行と混同しがちな女の子

「わたし(私)は、おひめさまが好き！」

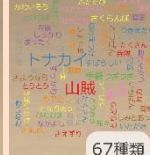
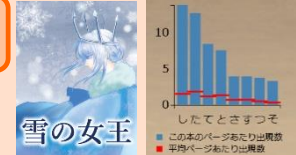
サ行タ行の発音を含む3音以上の言葉が多くて、30ページくらいで...

両者からの複合的な条件とよく一致する絵本を出力



言語聴覚士(構音訓練を行う専門家)

竹田綜合病院と共同研究中



関連文献

- [1] 服部正嗣, 小林哲生, 藤田早苗, 奥村優子, 青山一生, “ピタリエ: 興味・発達段階にピッタリの絵本を見つけます,” *NTT 技術ジャーナル*, Vol. 28, No. 6, pp. 54-59, 2016.
- [2] K. Aoyama, K. Saito, H. Sawada, N. Ueda, “Fast approximate similarity search based on degree-reduced neighborhood graphs,” in *Proc. The 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1055-1063, 2011.

連絡先

服部 正嗣 (Takashi Hattori) 協創情報研究部 知能創発環境研究グループ
Email: cs-liaison-ml at hco.ntt.co.jp



Innovative R&D by NTT
オープンハウス 2019