



第1回ナレッジグラフ推論チャレンジ2018

アイデア部門1

teamING

株式会社 数理先端技術研究所

2018/11/25 生島高裕



アジェンダ

1. 自己紹介
2. はじめに
3. 推理小説について必須情報
4. 推論方法考察
5. 将来の推理小説推論マシンの仕様
6. アイデア部門応募のための手作業リスト
7. 終わりに



自己紹介

ソフトウェアの会社で、いくつかの組織で活動 & 参加しています。

・全脳アーキテクチャ勉強会

<http://www.sig-agi.org/wba>

・LODI

<http://linkedopendata.jp/>

・サイバー絆研究所 (Institute for Cyber Associates, ICA)

<http://join-ica.org/ica/index.html>

・公共貨幣フォーラム

<http://www.muratopia.org/Yamaguchi/MoneyForum-j.html>

・AI社会論研究会

<http://aisocietymeeting.wixsite.com/ethics-of-ai>

・人工知能美学芸術研究会 (AI美芸研)

<https://www.facebook.com/groups/1155307841158207/>

・シンギュラリティサロン

<https://www.facebook.com/groups/SingularitySalon/>



はじめに


- 推論、推理などは人間の知性の最たるもの、これを AIで実装することは人類の長年の夢
- 根本は創造性で、手段として各種推論が出現？
- 推論とは
- 類推、帰納、演繹、アブダクション、抽象化、一般化、双対性が基本
 - 物理などのパラダムシフトを起こすための方法論
 - (相補性原理、対応原理、パラメータ問題、)なども範疇に
 - できればマーフィーの法則的ものも範疇に入れたい

この大きなテーマのもとまだらのひもの推論、推理をするに当たって考察
最終的にはプログラミング




推理小説について必須情報

- 伏線、トリック、アリバイ、動機、犯行手段、凶器、証拠などの推理小説常識
- 作り方がわかれば謎解きもできる。(ライバル作家同士の解析方法、メタ解析)
- 面白くなければならない(文学なのでエンターティメント性が必要、それも使える)
- 作家の設計視点(果たしてシステム設計と同じか?)
 - 起承転結のストーリー
 - 犯人構想
 - 動機の設定
 - わからないものがわかってくる楽しみ(謎解きの楽しみ)
 - 事件の発生(ホームズ登場前)
 - 解決への謎解きのストーリー(ホームズ開始)
 - 犯人との駆け引き(ホームズとの知恵比べ)
 - 推理以外の楽しみ(推理小説の肉付け)
 - 状況におけるうんちく
 - 事件の場所、時間における話題



推論方法考察

1. 根拠文からの推論
根拠に基づく推論(本格ミステリ:手がかりが全て示される)
手がかりだけから論理的に推論する方法
2. 状況文の推論
作者の意図、表現を読んだメタ推論(演出 [効果的に見せる]からの推論)
手がかりから推論するのではなく、文章中の微妙なニュアンスをくみ取り推論 (映像は楽)
作者の展開構造の知識から推論(設計視点)
3. 全作品からの推論
作者の全作品からの傾向分析に基づく推論
犯人の傾向、トリック傾向の知識ベースを作成し、それに基づき推論
「あの作者はこのような人物をよく犯人する傾向がある」などの統計分析及びデータマイニングによる推論



将来の推理小説推論マシンの仕様

今回は人間の手作業

その手順をそのままアルゴリズム化し、自然言語処理、ナレッジグラフ、推論エンジンなどのツールを使用する

以下は1パスの例、今回の人間による実験では5パス

mainプログラム

1. while(readSentence) { #一文読込
 2. if (judgeSubjectSentence) #主題に関係する文か、推論に関係ない記述かを判定
 3. parseSubjectSentence #主題に関する文の処理
 4. if (judgeConcludedSentence) #結論文についての記述チェック
 5. parseConcludedSentence #結論に関する文の処理
 6. if (EOF) exit #終了判定
- }



将来の推理小説推論マシンの仕様(主な関数)

1. judgeSubjectSentence(主題に関係する文か、推論に関係ない記述かを判定)

基本的にはターンテイキング問題、フィラー発見と同じ

1パス:根拠文、状況文(斜字)を意識して抽出する

2パス:推論情報抜出実行、関係ない文は削除

2パス:重要語彙関連は太字にする

3パス:根拠文、状況文(斜字)の重要語彙関連(太字)太字のみを抽出

2. parseSubjectSentence(主題に関する文の処理)

2パス:推理小説のフェーズ挿入、以下フェーズ例、フェーズ文作成は全体の意味把握が必要

3パス:推理小説の7つのフェーズに対して推理内容のラベルづけの実行、「要約からの推論」の実行



将来の推理小説推論マシンの仕様(主な関数)

3. judgeconcludedSentence (結論文についての記述チェック)

結論文の抜粋実行

4. parseConcludedSentence (結論に関する文の処理)


4パス: 推理小説の7つのフェーズに対して直接説明できる最小文章の抜粋とその説明最小文章要約の追加

5パス: 推理小説の7つのフェーズ、推理内容のラベル、要約からの推論、最小文章要約のみをまとめる。
「結論と説明」が最終アウトプット
自推理との比較



アイデア部門応募のための手作業リスト(1パス、2パス)

1. まだらのひも(原文からの推論情報抽出[根拠文、状況文])
根拠文、状況文(斜字)を意識して抽出する
2. まだらのひも(推論情報抜粋部に対し重要語彙関連抽出とフェーズアノテーション実行)
推論情報抽出実行、関係ない文は削除 (p20-p14に絞込)
重要語彙関連は太字にする
推理小説のフェーズ挿入、以下フェーズ例、
フィルター的な認識が必要、フェーズ文作成は要約、全体の意味把握が必要
<No1 初期立ち上げ、背景、状況説明> [7つのフェーズ]
<No2 状況説明を聞いた後のホームズの疑問点、確認事項の質問、事件に対する仮説生成は始まっている>
<No3 大仮説の裏とり、動機の確認>
<No4 ここから大仮説の検証過程、細部のトリックの解明がスタート>
<No5 ホームズはトリックをこの時点で解明、ここでトリックの解明ができればベスト>
<No6 トリックの結論、ここまでトリックの解明が必要>
<No7 解説部>




アイデア部門応募のための手作業リスト(3パス)

3. まだらのひも(重要語彙関連抜粋と各フェーズに対する推理内容のラベル付の実行)
根拠文、状況文(斜字)の重要語彙関連(太字)太字のみを抽出(p14-p11に絞込)
推理小説の7つのフェーズに対して推理内容のラベルづけの実行、「要約からの推論」の実行例)
＜No1 初期立ち上げ、背景、状況説明＞ [7つのフェーズ]
【大まかなテーマの提示】 [推理内容のラベル]
【ジュリアの死に至る説明]
★要約からの推論 [要約からの推論] ここまでで推論帰結
この時点で説明の多さから犯人はロイロット博士と推定できる
(作者の小説構成からの推論)
対抗馬はロマの一団ということも候補に上がる
(今後の展開での記述量で推定できる:文章量が多いことはより犯人であることの可能性が高いという仮説に基づく)

アイデア部門応募のための手作業リスト(4パス)

4. まだらのひも(直接説明できる最小文章の抜粋とその説明) **ここから説明性**
推理小説の7つのフェーズに対して直接説明できる最小文章の抜粋とその説明 **最小文章要約**
の追加 (抜粋と追加でp11変化なし)
例)
<No1 初期立ち上げ、背景、状況説明> [7つのフェーズ]
【大まかなテーマの提示】[推理内容のラベル]
グリムズビ・ロイロット博士の死について風聞が広がっていること [最小文章の抜粋]
●最初にロイロット博士が死んだ話がある(犯人候補データ1)[最小文章要約]
「非常に興味深い特徴があるようです——飾りだけの引き綱、風の入らない通風口。
よろしいですね、ストーナさん、もうひとつ内側の部屋を調べに参りましょう。[最小文章の抜粋]
●穴と引き綱、通風口、完全密室情報しかし穴はある(トリックに関する部屋データ)[最小文章要約]



アイデア部門応募のための手作業リスト(5パス)

5. まだらのひも(推論情報のみ)

推理小説の7つのフェーズ、推理内容のラベル、要約からの推論、最小文章要約のみを**まとめ**る。

「結論と説明」が最終アウトプット(総量p4)

例)

＜No3 大仮説の裏とり、動機の確認＞ [7つのフェーズ]

【ホームズの書類調査結果】[推理内容のラベル]

●ロイロット博士の娘が結婚すると受取額が減るとい**う**動機確認(犯人候補データ10)[最小文章要約]

★要約からの推論

動機の調査、ロイロット博士には動機があるが、ロマには動機不足の感がある



アイデア部門応募のための手作業リスト(5パス:結論と説明)

犯人:ロイロット博士

直接の説明:No5までの動機、推定、トリックの解説が説明になる。

間接の説明:容疑者はロイロット博士とロマー団しか存在しないので犯人候補データの数でロイロット博士が推定できる。また状況のメタデータで犯人は推定できる。

トリック:

1. 寝室の配置はロイロット博士とジュリア、ストーンと隣
2. ジュリア、ストーンは完全密室
3. 完全密室であるがロイロット博士の部屋と引き綱、通風口の小さな穴で繋がっている
4. 口笛と金属音でまだらのひもを穴から侵入させ、寝台固定で動けなくして死亡させる
まだらの紐から蛇は難しい。しかし蛇からまだらの紐は簡単。(人間でも難しい)

感想:

犯人の推定はすぐできる。

トリックについても「本格ミステリ」の場合なので推論は可能と思われる。

人間によるシミュレーションでは推定できた。



アイデア部門応募のための手作業リスト(初期課題)

課題:

フェーズの分解が内容の変化で判定できるか？(人間はできる:意味がわかるから)

推理内容のラベルの作成(要約)ができるか？(人間はできる:意味がわかるから)

(しかし推論には不要かもしれないが、説明付与については必要と思われる)

要約からの推論はかなり難しい(人間でも人それぞれの表現になる)

最小文章要約については語彙集を準備してパターンを作っておけば可能かもしれない

この作業自体高度な作業になる(語彙集とパターンの自動生成になるので)

説明文の生成は上記の作業でオントロジー構造を作成しておかなければならない

その後はシリアライズすることにより説明文の生成ができる(人間はこれをやっている)

ただ、どの切り口で説明開始するかは説明方法のメタデータ生成が必要になる

(これも人間の場合できる人とできない人がいる)



最後に

課題(2回目の考察での)

コンパイラ方式なので内部状態の定義が必要(人間の場合それがオントロジー構造?)

推理小説の場合はアートの中でもかなり大衆的なので内部状態構造がシンプルなはず?!

そのデザインができればパズルのたびに状態変数をアップデートできる。(人間はやっている)

内部状態構造自体は人間の場合、成長と共に変化する(遺伝子と環境、先天性と後天性)

この構造がなければ推理小説の推論もできない。

そして必要なのは推理小説の知識、「アリバイ」、「裏を取る」などの推理小説における常識

人間は常識を必要に応じて取得する。その構造が推論マシンに必要。

不完全知識に対して補充しようと言うモチベーションがなければならない。


その機能を人間は「自我」で持っている。最終的には推論マシンにも自我が必要であろう。

そうして、仮説生成、創造行為が必要になる。

最近のアブダクションはパースの原点回帰しているようである。創造と区別がつかない。。

類推、演繹、後ろ向き推論は新しい概念を作るわけではない。

しかし、帰納、アブダクション、抽象化、一般化、双対性では新しい概念を生成する。



ご清聴ありがとうございました