

記事種別 レポート レビュー ハウツー インタビュー 連載 コラム

ニューストップ > テクノロジー > サイエンス

【コラム】  
世界を彩る 数学レシピ

16 テキスト解析×数学

日本数学検定協会  
[2014/03/14]

今日は富士ゼロックス株式会社でお話をうかがいます。富士ゼロックスではお客様の声を経営に生かすために、「意味を数える」システムが使われているということです。「意味を数える」とはどのようなことなのでしょう。富士ゼロックス株式会社・研究技術開発本部の舘野昌一さんにお話をうかがいました。



富士ゼロックス R&D スクエア

-本日はよろしくお願いします。今日はこちらで「意味を数える」ということが行われていると聞いてやってきました。

はい、お客様の声をより経営に反映させるために、「意味を数える」システムを作っています。そのことをお話しましょう。

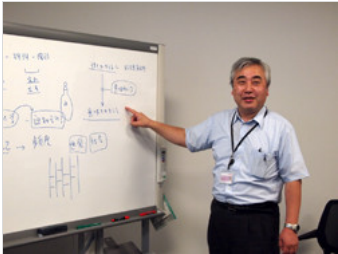
-お客様の声というと、たとえばインタビューやレビューといったものですか？

そうですね。インターネットの評判サイトなどに書かれている書き込みをイメージしていただくといいでしょう。お客様の声をどれだけきちんとくみとれるかは、経営の方向性を決めるうえでとても重要なものです。私はもともと慶応義塾大学の湘南藤沢キャンパスで、ソシオセマンティクスの第一人者である、深谷昌弘教授と一緒に活動していました。私も准教授として教えていたのですが、このソシオセマンティクスを応用して、お客様の声をくみとるシステムを作っています。

-「ソシオセマンティクス」とはどのようなものなのでしょう？

和訳すると「社会意味論」ですね。人それぞれによって、ものごとの意味のとらえ方は異なりますが、これは人がそれぞれでものごとを意味づけしているからです。人びとはこのような意味づけをもとに社会現象を作り出していると言えるわけですから、社会現象を研究するには、この意味づけを研究しなければなりません。そういった思いから、深谷昌弘教授がこの学問を作りました。ソシオセマンティクスを応用して、私たちは「意味を数える」ということを試みています。

-わかったような、わからないような…難しいですね。ソシオセマンティクスを応用して、具体的にどのような分析が行われているのでしょうか？



具体的にどんなことを行っているのかをご説明します。まず、分析するという行為は「定量分析」と「定性分析」にわけることができます。液体を例に考えると、「定量分析」は液体の量を知るための分析、「定性分析」は液体の性質を知るための分析です。どちらが難しいかというと「定性分析」のほうが難しいですね。

-そうですね。液体の量はすぐに測れますけど、液体の性質はリトマス試験紙などを使わなければ分かりませんか  
らね。

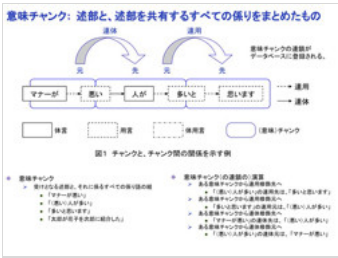
そういうことです。この分析を情報にあてはめたらどうなるでしょうか。情報分析も定量、定性に分けられるの  
です。お客様からいただいたアンケートの文章があるとなると、定量情報とはお客様の性別や年齢などです。し  
かし、定性情報となると難しいですね。

-定性情報は文章の意味ということですよ…たしかに難しいですね。

そうです。だからこの意味を定量分析ができるような形にする、これが「意味を数える」ということです。

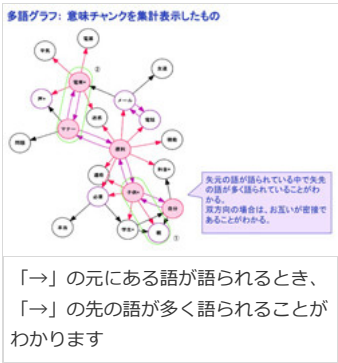
-なるほど。数えられるかたちにすれば、分析もしやすいですからね。

まず、文章内の語を数えるのは簡単ですよ。 「歩きます」なら「歩く」＋「ます」のように活用をとりのぞき  
ます。これを「形態素解析」と言います。次に連用や連体といった、係りをすべてまとめていきます。これを  
「意味チャンク」といいます。



-なるほど、係りによって意味が生まれているわけですからね。これを数えられるかたちにするわけですか。

たとえば、携帯電話の不満についてのアンケートがあったとします。このアンケートの「意味チャンク」を集計  
してできあがるのがこのような図です。



-なるほど、確かに「意味を数える」ということができるようになりましたね!

ただ、忘れてはならないことがあります。意味というのは言語の中にあるのではなく、人の頭の中にあるとい  
うことです。たとえば「美しい少女の母」という言葉があったとします。さて、美しいのは少女と母、どちらで  
しょうか?

-どちらの意味でもとらえることができますね…。この文章だけではどちらかを判断することはできないと思いま  
す。

はい、美しいのが少女なのか母なのか、言語のデータだけでは意味が確定しません。これを判断する方法が1つだけあります。わかりますか？

**-判断する方法ですか…。難しいですね、わかりません。**

答えは実際に見ることです。結局実際に見ないと美しいのは母なのか娘なのかわからないですよ。これはコンピュータでは解決できません。

このような現場の情報も含めて言語を分析しようという学問、これを認知言語学と言います。意味は言語に存在せず、人の頭の中に存在するわけです。だから、この意味そのものをコンピュータで置き換えることはできませんよね。

**-お客様の声を経営に生かすためのシステムは、理系的なアプローチと文系的なアプローチ、どちらも必要というわけですか。**

そうですね。「意味チャンク」もコンピュータの支援を得ながら人間が楽をするというシステムをめざして集計されています。コンピュータが解析をして、人間が解釈をする。この両方の作業がスムーズに行われる前提として、ツール作りをしているんです。

**-コンピュータと人間、それぞれの役割を意識したシステムなんですね。ところで舘野さんは、学生時代にどんな勉強をされていたんですか？**

中学生のころは受験勉強よりも秋葉原通いがおもしろく、ジャンクのPCパーツなどを買い集めては、計算機などを作っていました。

ほかに、小学6年生と中学2年生のときに科学センターに行き、理科好きが確定しましたね。「頭の体操」という本も好きで読んでいて、数学はクイズみたいでおもしろかったです。

**-数学はお好きでしたか？**

数学は中学まで得意でした。高校に入っても行列などはとてもおもしろくて、あみだくじについて考えたことをレポートにしたりしていました。

あとは、プログラミング問題によく取り組んでいましたね。

**-ありがとうございます。本日は貴重なお話をありがとうございました。**

今回、舘野さんのお話を聞いて数学の文章問題のことを考えました。数式をどれだけ解けたとしても、問題文という言語を読み解けなければ式をたてることはできません。小学生の頃は「問題文をきちんと読みましょう」とよく言われましたが、数学を活用するには国語的なことも大事なのですね。このように学際的な思考を持つことで、数学の世界もさらに味わい深くなる気がしました。舘野さん、貴重なお話をありがとうございました！



#### 今回のインタビュー

##### 館野 昌一(たての まさかず)

慶應義塾大学大学院工学研究科管理工学専攻修士課程修了。富士ゼロックス株式会社入社後、おもにシステム製品の計画に従事。

米国ゼロックスで計算言語学(日本語処理)の研究を行った後に帰国。慶應義塾大学・深谷昌弘教授との共同研究によりテキスト意味空間分析法の確立へ向けて研究とソフト開発を進める。

現在、人とコンピュータが協力して、人々の思いを集約し、よりよく伝えることをめざし、研究とソフト開発に従事。テキスト意味空間分析検定の創設に意欲。



このテキストは、(公財)日本数学検定協会の運営する数学検定ファンサイトの「数学探偵が行く!」のコンテンツを再編集したものです。

(公財)日本数学検定協会とは、数学の実力を図る数学・算数検定の実施をはじめ、数学に関する研究や講習会、普及、啓発活動を通じて数学的な思考力と応用力を提案する協会です。

数学検定・算数検定ファンサイトは、(公財)日本数学検定協会が開設する、数学好きが集うコミュニティサイトです。社会と数学との接点や、数学教育に携わる人々の情報交換など、数学を軸により多くの示唆と教養を社会に提案していくために生まれました。本コラムのベースとなった「数学探偵が行く!」をはじめ、数学を楽しみながら学べるコンテンツをたくさん運営中です。

総合数理×数学

※本記事は掲載時点の情報であり、最新のものとは異なる場合があります。予めご了承ください。

インデックス

連載目次

- 第42回 [青果×数学](#)
- 第41回 [フォント×数学](#)
- 第40回 [睡眠×数学](#)
- 第39回 [ネットオークション×数学](#)
- 第38回 [オーガニック製品×数学](#)

[もっと見る](#)