

請求項と実施例との対応付けに関する考察

東京工業大学精密工学研究所客員教授／株式会社日立製作所 中央研究所 岩山 真

PROFILE

1992年株式会社日立製作所入社。文書検索、自然言語処理等の研究に従事。また、NTCIRにおいて特許検索用テストコレクションの作成に携わる。2009年度より特許版産業日本語委員会委員。

✉ makoto.iwayama.nw@hitachi.com

☎ 042-323-1111

1 はじめに

特許明細書は長いと読解に時間を要する。たとえば、特許の審査業務では、まず審査対象の明細書（本願）の内容を理解せねばならない。本願の権利は明細書の「請求項」に書かれているため、ここでは請求項の内容を理解することになる。しかし、請求項は、発明の権利範囲を広げるために抽象的になりやすく、請求項を読むだけで発明の具体的な内容まで読み取ることは難しい。一方、請求項の内容は、「発明の詳細な説明」（以降、本文と呼ぶ）に詳しく記述されているため、請求項のみで理解が不十分な場合は、対応する本文を読めば良い。しかし、一般に本文は長く、必ずしも請求項と対比した構造で書かれているわけではないため、請求項に対応する箇所を見つけるだけでそれなりの時間を要してしまう。

請求項に対応する本文の箇所を即座に特定することができれば、本願理解の効率も上がるであろう。本願理解のみならず、任意の明細書の読解にも役立つ。たとえば、検索結果の明細書を理解する際も、要約情報だけで不十分な場合は、請求項に対応する本文を選んで読むこともできる。

本節では、特許版・産業日本語委員会（特許オントロジー検討小委員会）で検討している、請求項と本文との対応関係の定義について報告する。本定義に基づいて明細書を構造化しておけば、上記の読解支援にそのまま役立てることができる。また、請求項に対応する本文の場所を（半）自動的に見つけようと試みる際にも、あらかじめ構造化されている明細書は訓練データとして役立つ。

2 懸案事項

請求項と本文との対応関係を定義するには以下の3つの懸案事項がある。

（1）請求項側の単位

請求項もそれなりに長いと、各請求項全体に対して対応関係を定義するのか、それとも請求項を何らかの基準で分割して、分割した単位それぞれで独立に対応関係を定義するのか、選択の余地がある。

（2）本文側の最小単位

明細書には段落（【n n n n】）という単位があり、対応先としてもこの段落を最小単位として用いるのが適切である。しかし、場合によっては、段落が長く、段落内の文や句といった細かい単位を用いた方が良いかもしれない。

（3）対応関係の種類

請求項と本文との対応関係も、細かく見れば様々な意味で対応している。どの粒度の関係を定義すべきか決めておく必要がある。

本稿では、上記の3つの懸案事項について検討した結果を報告する。検討には、東京工業大学において人手（非専門家）で構造化（対応付け）を行った明細書（インクジェットプリンタに関する98件の明細書）を用いた。

3 請求項側の単位

請求項に関しては、既に定義されているオントロジー（「図式産業日本語・検索産業日本語（平成22年）」1.1節）に基づいて検討する。図1に例を示す。なお、以降の例で表示している明細書の一部は、全て特許電子図書館（IPDL）からの引用である。

(98003851)

【請求項1】基板上の対向する電極間に、導電性薄膜形成用材料を含む溶液の液滴を付与して導電性薄膜を形成し、該導電性薄膜の一部に電子放出部を形成する電子放出素子の製造方法であって、前記液滴の付与を湿度が70%以下に保持された雰囲気下で行うことを特徴とする電子放出素子の製造方法。

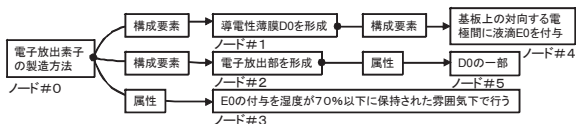


図1：請求項の構造化例

ここで問題となるのは、請求項のどの粒度のノードに対して本文との対応関係を定義するかである。全てのノードに対して対応先の本文が同定できればそれに越したことはないが、コストがかかる。本節では、一番粗い粒度である請求項全体（図ではノード#0）と、請求項全体を分割した単位（図ではノード#1、ノード#2、ノード#3）とを比較検討する。

(98003851)

【請求項1】基板上の対向する電極間に、導電性薄膜形成用材料を含む溶液の液滴を付与して導電性薄膜を形成し、該導電性薄膜の一部に電子放出部を形成する電子放出素子の製造方法であって、前記液滴の付与を湿度が70%以下に保持された雰囲気下で行うことを特徴とする電子放出素子の製造方法。

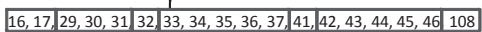


図2：請求項と本文との対応例

図1の請求項に対する対応先のデータを図2に示す。対応先の番号は段落番号である。たとえば、最初の構成要素（図1ではノード#1に相当）に対しては、【0029】～【0031】の3段落と、【0041】の1段落が対応している。この構成要素（ノード#1）の内容を調べたければ、4段落を読めば十分で、選択的読

みという点では適切な粒度であることが予想できる。更に細かいノードの対応先を同定させると、今度は、対応先の本文側を段落よりも細かくする必要が出てくる（たとえば、文レベル）。

請求項全体（図1のノード#0）に対して対応を同定する場合はどうであろうか。一つの請求項全体のトピックは広いと、対応する本文側の範囲も広くなってしまう可能性がある。図2を見ると、請求項全体に対しては、対応する段落数が18個にも及ぶ。請求項が1つしかない明細書の場合、明細書のほぼ全段落がその請求項に対応してしまうケースもあり、選択的読みという観点では適切ではないことが多い。

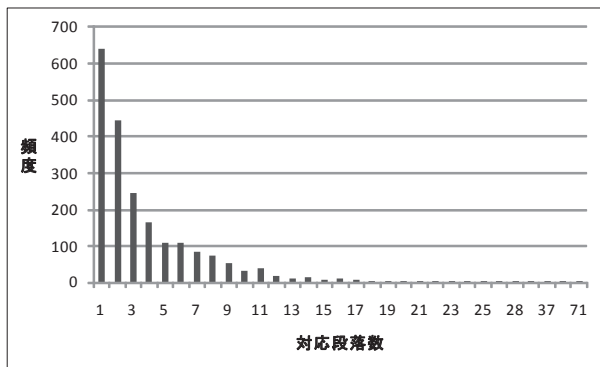


図3：対応段落数の分布

図3に、請求項の構成要素に対応している段落数のヒストグラムを示す。約70%の構成要素は、対応する段落数が5個以下となっている。このことから、請求項側の単位としては、請求項全体に対応するノードの直下のノードが適切であることが予想できる。

4 本文側の最小単位

本文側の粒度は、請求項側の粒度に合わせる必要がある。請求項に対応する本文を読むという設定では、目的の請求項に関係のない本文までは読みたくない。よって、本文側の粒度は、請求項側の粒度と同等かそれより細かいことが望ましい。

前節での考察から、請求項側の粒度は、請求項全体を

大きく分割する構成要素が適切そうであることがわかった。よって、本文側の粒度の候補である段落は、請求項側の構成要素と同等かそれより粒度が細かければ良い。本文側の粒度が請求項側の粒度より粗くなってしまうケースとは、同一の対応先（段落）に対して、複数の異なる対応元（請求項の構成要素）があるケースである。実際にデータを見てみると、このようなケースはほとんど見られなかった。図4に事例を示す。

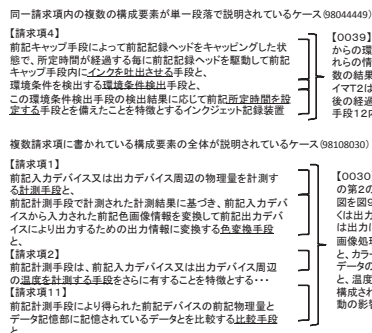


図4：多対一の対応例

最初の事例は、同一請求項の複数の構成要素が単一の段落に対応している場合であり、請求項全体を1つの段落で説明している。次の事例は、異なる請求項の構成要素が単一の段落に対応している場合であり、請求項の構成要素と図面の番号との対応を記している。両者も対応元の粒度より対応先の粒度が粗いが、そもそも対応先となる段落の長さが極端に長いわけではないので、無駄な部分を読むとは言え、その範囲は狭い。

問題は、対応先の段落が長すぎて肝心の対応部分がピンポイントで見つけられない場合である。今回は検討できなかったが、そのようなケースが多い場合は、段落を分割するなどの処置を施す必要が出てくるかもしれない。

上記のような問題の可能性は残るものの、ここでは、対応先の本文の粒度としては段落（【n n n n】）が適切であると結論する。

5 対応関係の種類

請求項と本文との対応関係のほとんどは、請求項の内容を本文で詳細に説明しているケースであった。この他、請求項の構成要素の効果を述べているケースが散見されたが数は少ない。図5に事例を示す。効果を記載している部分に下線を引いている。

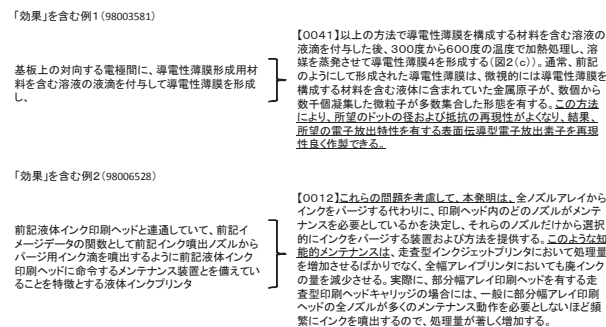


図5：「効果」を含む対応関係の例

これらの2件では、効果だけではなく詳細な説明も合わせてなされている。検討に用いたデータからは、対応先の段落に効果のみが記載された例は見つからなかった。よって、「詳細化」以外にも「効果」という関係はあるものの、区別して別扱いするまでもないと考え、今回の検討では、「効果」もあわせて「詳細化」として扱うことにした。ただし、今回のインクジェット・プリンタ以外の分野に関して検討を加える必要がある。

6 まとめ

以上の検討結果を踏まえ、請求項と本文と対応関係を、現状で以下のように定義する。

(1) 請求項側の単位

請求項を大きく分割する構成要素とする。

(2) 本文側の最小単位

段落（【nnnn】）とする。

(3) 対応関係の種類

細かく場合分けせず、対応の有無の2値とする。

最後に例として、図1の請求項の構造に対応関係を付記したものを図6に示す。

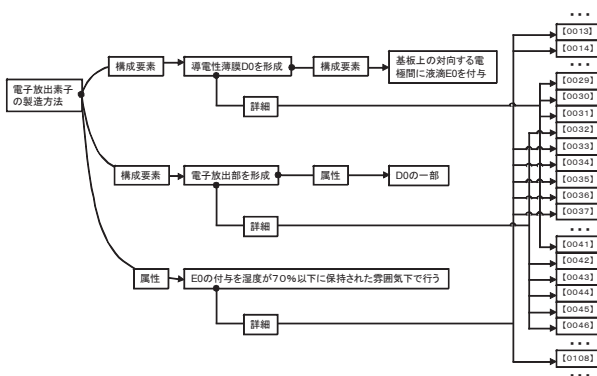


図6：明細書全体の構造化例

参考文献

- ・「2010（平成22）年度特許版・産業日本語委員会報告書」
<http://www.japio.or.jp/kenkyu/files/kenkyu04/ptj-h22.pdf>