

物性間の関係性データベース作成（下）：マテリアルキュレーション®のために

Construction of Database on Relations among Physical Properties (2)

物材機構¹ °吉武 道子¹

NIMS¹, °Michiko Yoshitake¹

E-mail: yoshitake.michiko@nims.go.jp

科学原理も含めた材料情報を分野横断的に活用して材料探索を行うために、物性間の関係（科学法則の知識）をデータベース化し、その関係性を探索するシステムの開発を行っている[1]。今までは、発表者が場合に応じて文献調査を行い、物性間の関係性を手動で見つけてきた。しかし、システムとして広く利用できるようにするには、物性間の関係性データベースが必要で、その作成ための何らかのコンピュータ支援技術が必要である。一つ前の発表、「物性間の関係性データベース作成（上）」では、まず、既に関係性が有るとわかっている性質を表現する術語同士の間、自然言語処理的に関係性を見つけることが出来るかどうかを調査した。

本報告では、既に関係性が有るとわかっている性質を表現する術語ペアを学習データとしてコンピュータに入力し深層学習させて学習済みモデルを得、この学習モデルを用いて、元の文書からどの程度、関係性が有るという箇所を抽出できるかについて検討した。評価は、表1に示したA、B、Cで表現される領域にどのぐらいのデータが含まれるかにより行った。まず、学習データに含まれる関係性が全て抽出できるか

表1 処理結果の評価のための区分

(出来ないものを drop)、

$$drop = \frac{C}{A + C}$$

どれだけ無関係なデータを抽出したか
(noise)

$$noise = \frac{B}{A + B}$$

		学習データ	
		含まれる	含まれない
出力データ	含まれる	○(A)	幅広(B)
	含まれない	(C)	○

で評価した。

また、noise にも種類があり、種類別の検討を行った。当日の講演では、学習データをどのように作成したか、学習モデルを用いて抽出することによって学習データ作成時に混入していた誤りが発見できた場合などに言及する予定である。

[1]吉武道子,桑島功,柳生進二郎,知京豊裕：表面と真空, 61, 200(2018).