金属材料論文の知識メディア化の試み

An Application of Knowledge Medium to semi-automated summarizing of Technical Papers in Metallurgy

松尾利行† 武田英明 西田豊明

Toshiyuki Matsuo Hideaki Takeda Toyoaki Nishida

奈良先端科学技術大学院大学 情報科学研究科

Graduate School of Information Science, Nara Institute of Science and Technology

In this paper, we apply Knowledge Medium System(KMS) to semi-automated summarizing of technical papers in metallurgy. We propose Knowledge Piece(KP) is a very powerful and robust method, because it is a knowledge chunk which has an appropriate grain size to textual process. KP consists of several sentences and linkage attributes toward other KPs. We develop METIS(METallurgy papers Intelligent Surveyor) system which acquires KPs from technical papers in metallurgy, summarizes each paper and survey plural papers from KPs. We estimate METIS performance using ten technical papers with experts.

1 はじめに

金属材料技術研究所では、新しい金属材料の開発あるいは既存の金属にある種の操作を加えることで金属特性の向上をするための研究を行なっている。金属材料に関する研究論文は数多く発表されており、今回対象とする超合金に関する論文も、年間１００~２００件程度発表され、しかも過去数年間に発表された論文を累積すると検索の対象となる論文は１０００件程度にまでのぼる。これら研究論文は、その内容を読解するのに多くの時間を要するだけでなく、一度読解された論文要約内容の表現形式が研究所内で統一化されていないなどの問題がある。

人工知能の立場から問題解決を試みるために、まず金属材料論文を実際にサーベイし、さらに金属材料の専門家からの指導を受けた。その結果、この種の論文がかなり構造化されており、しかも専門用語が随所に出現しているため、専門用語を使ったヒューリスティックなテキスト処理が可能であることが分かった。

そこで、ヒューリスティックなテキスト処理と知識メディア化 (Stefik[7]) の手法を用いて、プロトタイプシステム (METIS システム:METallurgy papers Intelligent Surveyor) を作成した。このシステムは、マークアップ言語 (マークアップ言語として、ＳＧＭＬを採用すべきであるが、本研究のプロトタイプシステムはＨＴＭＬ文書を使用している) で書かれている原著論文を入力とし、原著論文の要約および論文のサーベイを出力することができる。原著論文の要約としては、(1) 論文内容の構造化による要約 (図 1参照)、(2) 金属特性・要因間の因果関係の一覧表形式による表現 (図 2参照)、(3) 論文中の全図表のリンク結合による論文内容のマップ表現 (図 3参照) がある。ただし、(3) のみは現在まだ完全自動化はしていない。また、複数論文のサーベイとして、(4) 類似度による論文の空間配置 (図 4参照) と、(5) ３次元表示による因果関係の分布 (図 5参照) があり、これらはいずれも自動化されている。マップ表現とは、２次元座標上に知識片の集合を配置し、構成する知識片およびその結合状態で知識を表現している。また、空間配置とは、マップ表現と同様に２次元上に知識片を配置するが、配置位置情報により知識を表現する。

ＫＰ (KnowledgePiece: 知識片) という概念を導入することにより、ヒューリスティックな処理だけで知識メディア化が可能になった。ＫＰとは、弱い構造をもった知識メディア表現の１種であり、ＫＰの基本構造は、ボディ部、アンカー部、マッチング条件部からなる。ＫＰは、ボディ部はカードサイズ程度の情報量を持つ知識のチャンクであり、中間知識表現として適度な粒度である。

本研究では、プロトタイプを評価するために、１０件 (Khan84[8], Duhl71[3], Khan86[9], Baldan92[2], Caron83[6], Nathal85[5], Ge85[4], Hopgood91[1] など) の中で３件の論文で筆者がテストした後、残りの７件について知識ベースを無修正のままブラインドテストを実施し、金属材料の専門家から下記の評価を受けた。(1) 論文内容のマップ表現は、理解を深めるのに効果が大きいが、表示上の工夫が必要である、(2) 不適切な使用をしている専門用語がある、など多少の問題点はあるが、論文要約と論文サーベイは、文献の検索や理解を助けるのに効果が大きい。

2 知識メディア化について

2.1 金属材料論文の構造について

予備調査として、金属材料の専門家が作成していた各論文の要約を参考にし、さらに、専門家からの教授を受けた。金属材料の論文の多くは、(1)title, (2)authors, (3)affiliation, (5)literature, (6)date, (7)summarizer, (8)objective field, (8)purpose, (9)experimental procedure, (10)experimental result に構造化されていることが分かった。この構造の中では、(9)experimental procedure, (10)experimental result が重要であるため、(9) については、金属材料の名称及び組成、実験前準備、実験名称及び方法に、(10) については、実験結果、考察に細分化した。また、専門用語および英語の慣用句とのみマッチング処理を使うだけで、ほぼ分類できる見込みが得られた。

2.2 ＫＰ

本研究では、複数のＫＰ (KnowledgePiece: 知識片) を結合した論文情報ベースを構築した (図 6参照)。

ＫＰとは、(1) 人工情報系 (コンピュータ), (2) 自然情報系 (人間) のいずれにも理解可能であり、かつ、弱い構造を持つ知識メディア表現の１つである。ＫＰは、基本構造として、ボディ部、アンカー部とマッチング条件部から成る。

連絡先: 奈良先端科学技術大学院大学 情報科学研究科, 〒 630-01 生駒市高山町 8916-5, 電話:07437-9-9211(内 5316), fax:07437-2-5269,

(†日本総合研究所から出向中) E-mail: tosiy-m@is.aist-nara.ac.jp, URL: http:/ai-www.aist-nara.ac.jp/tosiy-m/

Improvement of Creep Strength in a Nickel-base Single-crystal Superalloy by Heat Treatment

BRIEF

1. super alloy
    1. CMSX-2
2. experiment
    1. creep
    2. tensile
3. figure & table
    1. creep
    2. deformation
    3. dislocation
    4. morophology
    5. γ
4. experimental result
    1. creep
        1. deformation
        2. strain
        3. strain rate
        4. behaviour
        5. primary creep
        6. secondary creep
        7. strength
        8. recovery
        9. test
        10. rate
5. discussion
    1. heat treatment
    2. chemical composition
    3. γ
    4. single crystal
    5. morphology
    6. columnar structure
    7. lattice
    8. misfit
    9. precipitate

図 1: 論文内容の構造化による要約

ボディ部は、テキスト情報や画像情報を扱う。テキスト情報として、ほぼカードサイズ程度の情報量すなわち数行のセンテンスの情報量を持つ。アンカー部は、他のＫＰへのポインターから成る。ＫＰ間の結合は、連想記憶である。すなわち、リンク情報はリンク先のＫＰ名のみとし、他のリンク属性は持たない。ＫＰを作成する基本的な手法が用語レベルでのマッチング処理である (後述) ので、マッチング条件部は、マッチングに使用する用語のリストからなる。

2.3 ＫＰの作成方法

ＫＰを作成するための基本的な手法は、次の通りである。

- 簡単な用語レベルでのマッチング処理であり、マッチングに成功した時、センテンス単位で情報を登録する。

Effect of processing conditions and heat treatments on mechanical properties of single-crystal superalloy CMSX-2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
creep life	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep deformation	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	○	–	–
creep elongation	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep strain	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep strain rate	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep behaviour	–	–	–	–	–	–	–	○	–	–	–	–	–	–	–	–	○	–
creep mode	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
primary behaviour	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
secondary behaviour	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep tertiary behaviour	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep strength	○	○	○	–	–	–	–	○	–	–	–	–	○	–	–	○	○	–
creep recovery	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep rupture	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep rupture life	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep test	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
creep rate	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	○	–	–
fatigue corrosion	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
fatigue limit								○									○	

図 2: 金属特性・要因間の因果関係の一覧表形式による表現

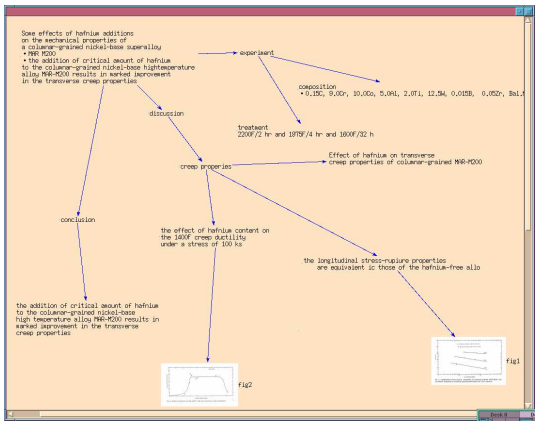

図 3: 論文中の全図表のリンク結合による論文内容のマップ表現

- 登録したセンテンス内に画像情報を参照していれば、画像情報をＫＰとしてリンクで結合する。

マッチングには、金属材料の専門用語辞書と英語慣用句辞書の 2 つの辞書を使用する。

1. 専門用語辞書
   専門用語を固定的用語 (「金属の特性」あるいは「実験方法」など) と流動的用語 (「金属名称」など) に分類する。固定的用語は、金属関係のシソーラスをベースに、専門家からのインタビューにより構築した。流動的用語は、実際の論文から構築した。
2. 英語慣用句辞書
   論文を分析した結果、因果関係を表現している知識が非常に重要であり、ＫＰ生成時、専門用語から知識を抽出するよりも、むしろ因果関係を表現する英語独自の言い回しを利用した方が過不足なく知識を抽出でき

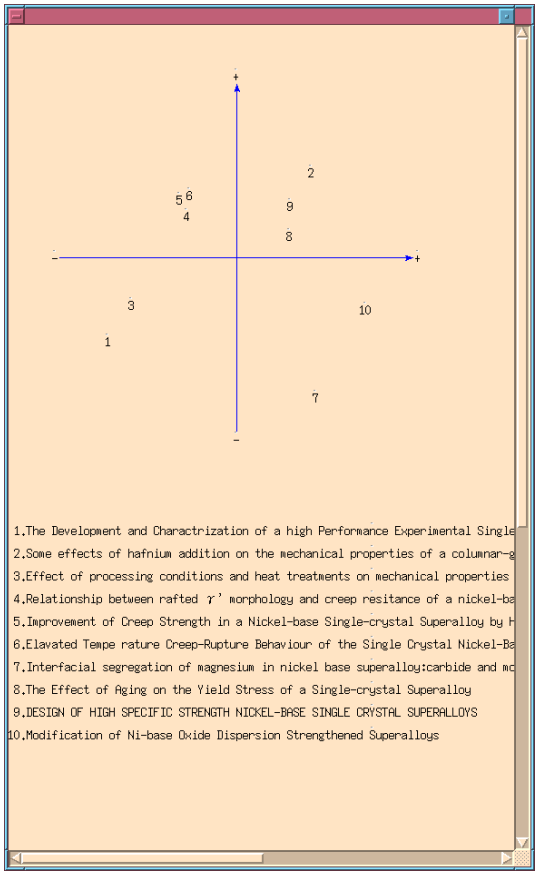

図 4: 類似度による論文の空間配置

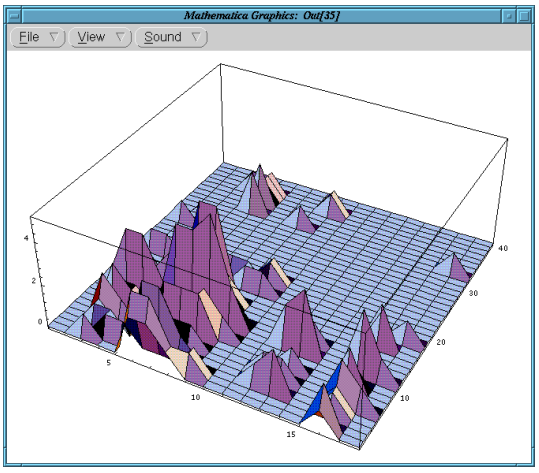

図 5: ３次元表示による因果関係の分布

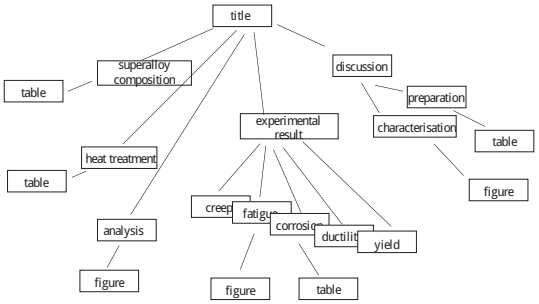

図 6: ＫＰを用いた論文情報ベース

ることが分かった。そこで、本研究では、２５パターンの言い回しを、因果関係を表現している英語慣用句辞書として構築した。

3　METIS システムの流れ

METIS システムは、マークアップ言語で記述した論文を入力とし、論文要約とサーベイを作成する。処理部は、(1) 論文からＫＰ作成部, (2) ＫＰから論文要約作成部, (3) ＫＰから論文サーベイ作成部から構成される (図 7参照)。(詳細は、次節参照)

4　論文からＫＰの作成

論文から各ＫＰを作成する方法を以下に示す。

1. superalloy composition, heat treatment, analysis, experiment, figure & table のＫＰについて
   各々のＫＰごとにマッチング条件部に、専門用語辞書の中から該当する用語を登録し、そのマッチングしたセンテンスを登録する。尚、 figure & table のＫＰは、一旦センテンスを抽出した後、実験結果あるいは観察結果別に分類する。
2. experiment result, discussion
   まず、因果関係を表現する英語の慣用句として、(1)affect, (2)arise, (3)be associated with, (4)cause など２５パターンを準備し、これらを含むセンテンスを抽出する。次に、金属特性別に分類して、 experiment result ＫＰに登録する。特性は、金属材料の専門用語の分類に従い、 (1)creep, (2)fatigue, (3)corrosion, (4)ductility, (5)yield に大分類した後、それら属性に応じ細目に分類した。また、要因別に分類して、 discussion ＫＰに登録する。要因は、 (1)preparation, (2)characterization に分類した後、細目に分類した。

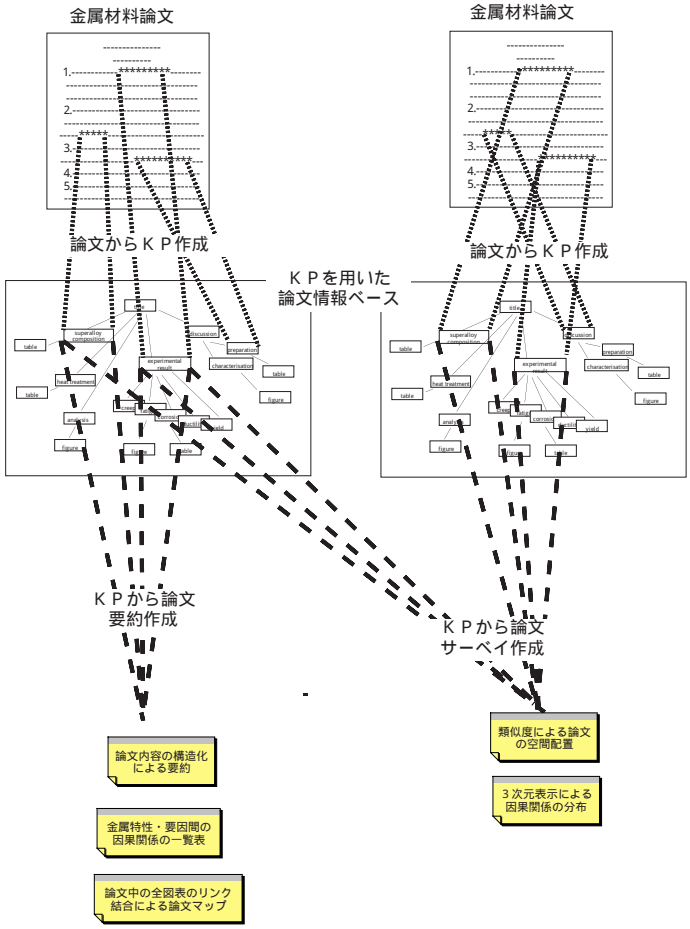

図 7: METIS システムの流れ

5　ＫＰからの論文要約作成

ＫＰ化された個々の論文情報ベースから、 (1) 箇条書, (2) 一覧表, (3) マップ の表現形式で論文内容を要約する。

1. 箇条書: 論文内容の構造化による要約
   各ＫＰに登録されている知識を、さらにキーワードまで要約し箇条書形式に要約する (図 1参照)。
   金属名称は、専門用語の中の流動的用語であり、本システムで未登録の金属を使用した論文には、対応できない。
2. 一覧表: 金属特性・要因間の因果関係の一覧表形式による表現
   論文要約 (図 1参照) では、明示されていない要因と特性の関連を表形式での表現している (図 2参照)。
   experimental result ＫＰと discussion ＫＰを用いて作成する。一覧表の縦軸が金属特性であり、横軸が各々の特性に関して考察している要因である。表形式で表現しているため、どの要因とどの特性について当該論文内で考察しているかを明示できる。現在のところ、(1) 複雑な因果関係 (例. 中間仮説があるような因果関係) が表現できない,(2) 因果関係の程度や関係の仕方までは抽出されない。
3. マップ: 論文中の全図表のリンク結合による論文内容のマップ表現
   論文内の図及び表をすべて抽出し、参照関係を用いて、関連あるＫＰから結合する (図 3参照)。
   論文構造の把握が容易になる。この部分はまだ完全自動化できていない。

6　ＫＰからの論文サーベイ作成

各論文のＫＰを統合して、 (1) 空間配置, (2) ３次元分布の表現形式で論文サーベイを作成する。

1. 空間配置: 類似度による論文の空間配置
   各論文の要約 (図 1参照) から、各項目間で類似度を

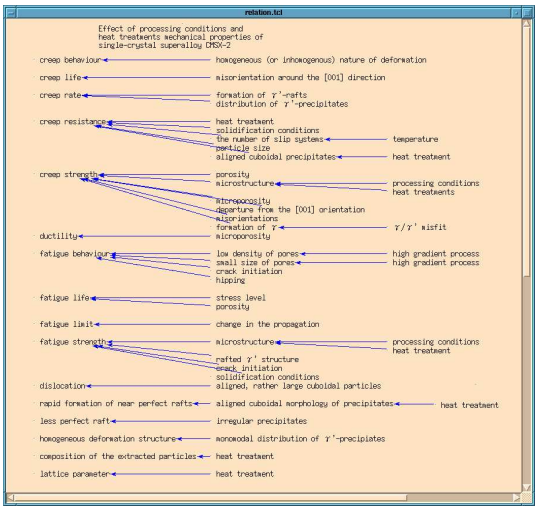

図 8: 因果関係ネットワーク

計算した後、それらを総和することで各論文間の類似度を計算する。さらに、求めた類似度に多次元尺度構成法を用い、２次元上に類似度を空間配置する (図 4参照)。
空間配置は、明示的に論文の類似度が示すので、類似論文の検索が容易できるが、現在のところ研究者の視点に応じた類似度の計算方法は実現していない。

2. ３次元表示による因果関係の分布
金属特性・要因間の因果関係の一覧表 (図 2参照) を論文方向に集計する (図 5参照)。
金属のどんな特性を向上することを目的とした論文が多いかまた、何を要因として考察している論文が多いかを明示的に示す。

7　評価

ヒューリスティックな知識は３件の論文から獲得し、METIS システムを実装した。知識ベースを無修正のまま７件の論文に適用できた。従って、数１０件あるいは数１００件の論文を入力としても、十分システムの頑強性を示すと思われる。たとえ、修正が必要な場合でも、ＫＰのマッチング条件部のみの修正で対応できる。

ＫＰからの論文要約の１つに、手作業で行っている因果関係ネットワーク (図 8参照) がある。因果関係ネットワークとは、一覧表 (図 2参照) では表現できない金属特性・要因間の複雑な因果関係をネットワーク形式で表現したものであり、因果関係ＫＰ (英語慣用句とのマッチング処理で自動作成する) に含まれるセンテンスを分析し、現在手作業で作成している。作成作業に長時間かかっているため、省力化する必要がある。

また、 METIS システムについて、金属材料の専門家から、下記の評価を得た。

「実用合金開発の分野では、直接必要な特性のみではなく、極めて多くの因子が関連してくる。これらに関する文献情報を的確に検索するのには長時間の作業が必要で、本研究の因果関係ネットワーク、論文類似度の表示は、文献内容の検索に役立つと思われる。さらに、どのような関係であるかを表示できるようになれば、意義は極めて大きい。論文内容のマップ表現は、理解を助けるのに効果が大きい。さらに工夫してほしい。ただ、不適切な使い方をしている専門用語があるので改善が必要である。例えば、 primary behaviour よりも primary creep の方が望ましい。 fatigue corrosion という用語は存在しない。 temperature ductility という用語は存在しないが、 high temperature ductility あるいは low temperature ductility である。 reduction だけでは、広い意味がある。 reduction in area の方が望ましい。」

8　まとめ

実際の金属材料論文に、中間知識表現としてＫＰを用いた論文要約およびサーベイする METIS システムを実装し、１０件の論文に適用した。また、論文要約およびサーベイについて、金属材料の専門家からの評価を受けた。

9　展望

システムが出力する中間知識表現であるＫＰを編集を可能にするＫＰエディタの開発中である。エディタでＫＰを修正することによって、専門家の要求や視点に応じた論文要約やサーベイ作成が可能になると思われる。

因果ネットワークの作成を支援するために、センテンス内のみの構文解析で、要因と結果を抽出できる可能性を検討している。すなわち、センテンスのパターンをいくつか準備し、そのパターンとのマッチング処理により、因果ネットワークの作成をしたい。因果ネットワークは、個々の論文の要約の理解を助けるだけでなく、複数の論文間で因果ネットワークを統合することで、新しい因果関係や論文間の矛盾を導く機能についても検討している。

本研究で導入したＫＰは、現在まだ単純な構造であるが、メタ知識の付加を検討している。ＫＰの組合せで、自動ドラフト作成システムにまで発展する可能性もある。

謝辞　金属材料の専門家である金属材料技術研究所の星本健一室長から、専門知識だけでなく、知識メディア化表現に関して有益なコメントをいただきましたこと、感謝致します。

参考文献

[1] A.A.HOPGOOD and J.W.MARTIN. The Effect of Aging on the Yield Stress of a Single-crystal Superalloy. *Materials Science and Technology*, Vol. 91, pp. 105–110, 1987.

[2] A.Baldan. RELATIONSHIP BETWEEN RAFTED γ' MORPHOLOGY AND CREEP RESISTANCE OF A NICKEL-BASED SUPERALLOY. *Superalloys*, pp. 647–656, 1992.

[3] D.N.Duhl and C.P.Sullivan. Some effects of hafnium additions on the mechanical properties of a columnar-grained nichel-base superalloy. *JOURNAL OF METALLS*, pp. 38–40, 1971.

[4] H.L.Ge, G.L.Chen W.V.Youdelis, and Q.Zhu. Interfacial segregation of magnesium in nickel base superalloy:carbide morphology and properties. *Materials Science and Technology*, Vol. 5, pp. 985–990, 1989.

[5] M.V.NATHAL and L.J.EBERT. Elevated Temperature Creep-Rupture Behavior of the Single Crystal Nickel-Base Superalloy NASAIR 100. *METALLUGICAL TRANSACTIONS A*, Vol. 16A, pp. 427–439, 1985.

[6] P.Caron and T.Khan. Improvement of Creep Strength in a Nickel-base Single-crystal Superalloy by Heat Treatment. *Materials Science and Engineering*, Vol. 61, pp. 173–184, 1983.

[7] Mark Stefik. The next knowledge medium. *AI Magagine*, Vol. 7, No. 1, pp. 34–46, 1986.

[8] T.Khan, P.Caron, and C.Duret. THE DEVELOPMENT AND CHARACTERIZATION OF A HIGH PERFORMANCE EXPERIMENTAL SINGLE CRYSTAL SUPERALLOY. *Superalloys*, pp. 145–155, 1984.

[9] T.Khan and P.Caron. Effect of processing conditions and heat treatments on mechanical properties of single-crystal superalloy CMSX-2. *Materials Science and Technology*, Vol. 2, pp. 486–492, 1986.