

目次

- ・ 2025年度学生と先輩の交流会 開催報告……長田孝二、林 潤 (pp. 2-3)
- ・ 2025年度中国四国支部 春の行事報告……豊島範男、佐藤重喜 (pp. 4-5)
- ・ series 研究最前線 (26) 文献情報を活用した物理モデル自動構築……加藤祥太 (pp. 6-13)
- ・ 第66回前期京岬会 (昭和33年卒業同期会) ……中村弥寿家 (pp. 14-15)

中国四国支部で2025年度春の行事が開催されました。
詳しくは記事をご覧ください。



企業見学会 (JR 四国高松運転所 : 2700 系特急形気動車の前で集合写真)

2025 年度学生と先輩との交流会 開催報告

学生会担当 長田孝二 (H4/1992卒)

林 潤 (H14/2002卒)

2025年6月28日（土）に京都リサーチパークにおいて「学生と先輩の交流会」が開催されました。京機会学生会（SMILE）が主催するイベントとして毎年開催されており、企業や公的機関で活躍されている卒業生と機械系の学部・大学院生との交流を深める貴重な場となっております。今年度は、64の企業や公的機関から82名の卒業生を含む126名の方々にご参加いただきました。

今回は例年の開催時期（12月頃）を半年前倒しして、6月末の開催となりました。前回開催日時と時期が近いいため開催に際する準備期間・場所など考慮すべき点は多くありましたが、SMILEメンバーおよび事務局の尽力、参加企業の皆様のご協力もあり盛会となりました。なお、開催期間・場所に関しては、引き続き、学生、参加企業様においてより良い形を引き続き模索していく予定です。

教室においては、前回に引き続き学部2年生、3年生への案内とともに修士1年生に対しても参加を積極的に呼びかけました。その結果、修士学生を中心におよそ150名の学生の皆さんが参加してくれました。

事前準備は前日の午後よりSMILEメンバーおよび事務局によって行れ、当日は下記のスケジュールの通りに滞りなく実施されました。

11：00 ～ 12：00 受付

12：10 ～ 12：25 京機会代表幹事挨拶（平方寛之さん（H9/1997卒））

交流会説明（SMILE代表 桑原和暉さん）

12：30 ～ 16：30 個別ブース交流会（1号館4F）

17：00 ～ 19：00 懇親会（1号館）

当日は好天に恵まれたこともあり、昨年度と比較しても開始直後から学生の参加者が多く、各ブースにおいて活発な交流がなされている様子が確認されました。また、機械系の多数の教員にもご参加いただき、運営にもご助力いただきました。



4F A会場（サイエンスホール）



B会場（G会議室）



交流会の様子

懇親会はブースのある1号館4階から1階のアトリウムに場所を移して開催されました。前回と同様に立食形式での懇親会であり、アルコールも提供されました。初めに参加者を代表して、SMILEのOBでもある豊田自動織機株式会社の勝野友介さん（H18/2006卒）にご挨拶と乾杯の音頭を頂きました。懇親会には、総勢170名近い参加があり、至る所で会話が弾んでいる様子がみられました。懇親会の終盤には、学生会担当教員（林）からのSMILEのメンバー紹介の後、SMILE代表の桑原和暉さんの一本締めで閉会となりました。

最後になりますが、交流会にご参加いただきました企業、公的機関の方々、ご準備いただきましたSMILEのメンバー（桑原和暉さん（M1）、森田未来心さん（M1）、児島はるかさん（M1）、杉野友紀さん（M1）、仁和一晟さん（B4）、納谷岳都さん（B4）、小島将吾さん（B4）、山本陽大さん（B4））、そして京機会事務局の段智子さんに深く感謝申し上げます。来年度もまた交流会で皆様にお会いできることを心待ちにしております。

中国四国支部 春の行事のご報告

中国四国支部長 豊嶋範男 (S57/1982卒)

中国四国事務局長 佐藤重喜 (H5/1993卒)

2025年5月17日(土)に春季行事を、昨年秋季行事に続いて高松で、19名の参加で開催しました。

集合時刻前には、希望された方でうどん屋に行き、本場のさぬきうどんを食べました。

企業見学会ではJR四国高松運転所を見学しました。カーブを高速で走行したときの遠心力を軽減するために車体を傾斜させる振り子式車両について、概要説明とともに特急南風号やうずしお号などに使用されている2700系特急型気動車を見学しました。

元本職や熱心なファンをはじめとした参加者の質疑は止まらず、懇親会へ持ち越しとなりました。

【講演会】

場所を移して総会後に講演会を開催しました。

大阪工業大学の貝原先生にお越しいただき、「生産システム×DX」と題してご講演いただきました。

単なる設計・生産のデジタル化ではなく、デジタル化により企業そして顧客の価値を生むことについて、企業と取り組んだ具体的な研究事例を交えて説明いただき、こちらの質疑も止まりませんでした。

懇親会では、企業見学会・講演会の延長戦をはじめ話は盛り上がり、あっという間のお開きでした。

秋季行事のアイデアもいただき、秋季も参加者に楽しんでもいただける行事を企画します。



講演会（大阪工業大学貝原先生）

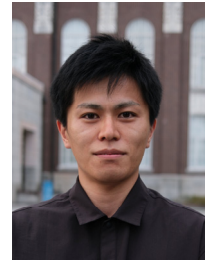


懇親会

series 研究最前線 (26)

文献情報を活用した物理モデル自動構築

加藤祥太 (H29 /2017 卒)



「複数の論文や書籍から物理モデルを作るAIを実現することはできるのか？」

1. はじめに

私は京都大学大学院情報学研究科情報学専攻人間機械共生系講座ヒューマンシステム論分野の助教として、自然言語処理技術を活用したプロセスインフォマティクスの研究に従事しています。学部は工業化学科（現在の理工化学科）、修士は化学工学専攻に所属していました。博士課程は情報学研究科に移りましたが、一貫して化学工学（実験室の化学を実生産に乗せ、その便利さを社会に引き出すための学問 [1]）の一分野であるプロセスシステム工学に関する研究に従事しています。本稿では、「文献からの物理モデル自動構築」の研究を取り上げます。

この研究で対象とする製造プロセスでは、費用や安全面の観点から、あらゆる条件において十分な数の実験データを確保することが困難な状況が存在します。例えば、化学プラントでは異常運転状態でのデータ取得は安全上のリスクを伴うため、正常運転時のデータしか利用できないことが多々あります。また、新規プロセスの設計段階では実機が存在しないため、実験データの取得自体が不可能です。このような状況でも、プロセスの挙動を表現し、運転最適化や異常検知を実現する必要があるため、質量保存則、エネルギー保存則、運動量保存則などの物理法則や現象論的な関係式などの科学的知見に基づいてプロセスの挙動を表現する物理モデルの構築が必須となります。

物理モデルの最大の利点は、データが存在しない状況におけるプロセスの挙動を表現できることです。しかし、その構築には対象プロセスに関する深い理解や専門知識が必要であり、関連する文献の調査を含めた多大な時間と労力を要します。私自身、博士課程在学時に、半導体原料である結晶製造プロセスを対象に物理モデルの改良に取り組みました [2]。しかし、対象プロセスの理解、100報以上の関連文献の精読と関連性整理、構築した物理モデルの性能評価と改善点の特定、

改善策の選定などが必要であり、数ヶ月の時間を要しました。

この物理モデル構築工程を効率化するため、文献情報から物理モデルを自動で構築する人工知能（Automated Physical Model Builder; AutoPMoB）の実現を目指した研究を進めています。

2. 人工知能による物理モデル自動構築

AutoPMoBは、人間が行っている物理モデル構築プロセスを計算機上で自動化することを目的としており、対象プロセスに関連する文献を自動収集し、それらの文献から必要な情報を抽出・統合して、目的に適合した物理モデル候補を自動生成するシステムです。Fig. 1に示すように、AutoPMoBは5つのステップから構成されます。まず、対象プロセスに関連する文献をデータベースから自動収集します。これは、人間が「蒸留 モデル」といったキーワードを用いてGoogle Scholarなどで検索する作業を、計算機が自動で行うというイメージです。次に、収集した文献の形式を統一します。一般的に用いられているPDF形式や画像形式の文書を計算機で処理可能な形式に変換します。続いて、文献中から変数や数式などのモデル構築に必要な情報を抽出します。例えば「The concentration of A, C_A , depends on the temperature T .」のような文から「 C_A : concentration of A, T : temperature」という情報を自動で読み取ります。その後、異なる文献から抽出した情報を整理し、同一の物理量や数式の表記を統一します。例えば、温度を表す変数が文献によって T や T_R のように異なって記載されている場合に、これらが同一の物理量であることを認識し、同じ記号で表現します。最後に、表記を統

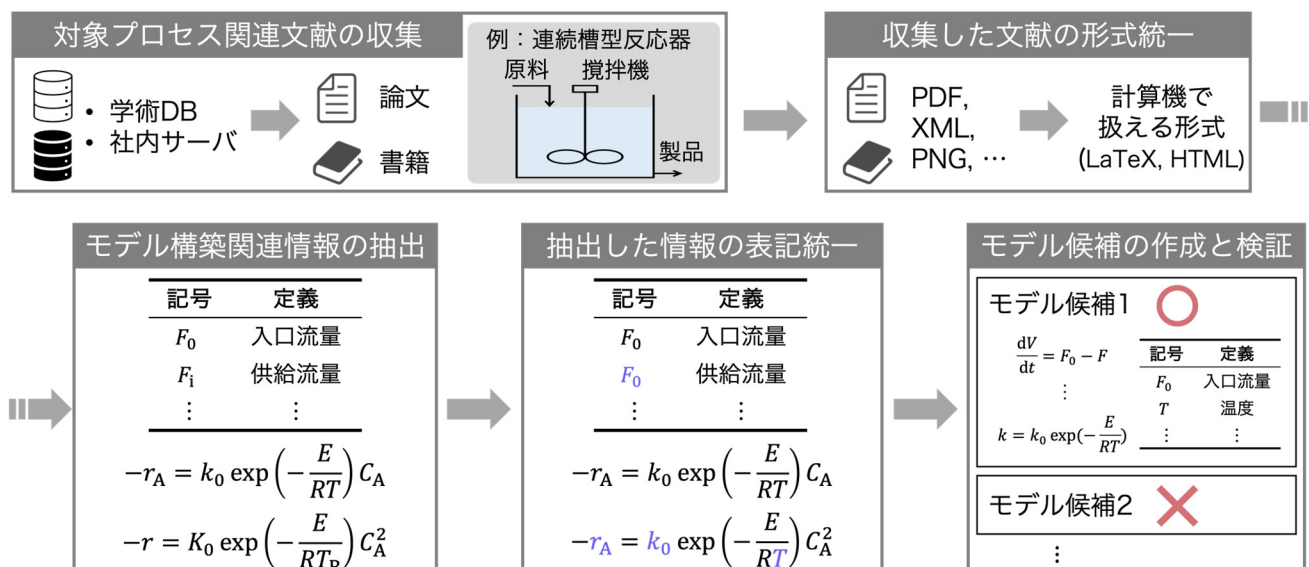


Fig. 1 物理モデル自動構築システムを構成する5つの処理

一した情報を組み合わせて、目的に合致した物理モデル候補を構築します。

これまでに、AutoPMoBの実現を目指して、情報抽出以降の要素技術を開発してきました [3, 11, 12, 15–17]。以降では、その技術の一部を説明します。

3. 論文からの情報抽出

物理モデルの構築には、変数や数式、実験データなど、複数の情報が必要です。その中でも、変数記号とその定義、および変数間の関係を表す数式は特に重要です。そのため、これまでに、変数記号・変数定義・数式を文書から正確に抽出する手法の開発に取り組んできました [3, 11, 12]。以降では、LaTeXやXMLといった文字列のみで構成される文書を対象とします。

LaTeXやXMLといった文書構造を表す形式の文書では、変数記号や数式は、斜体や専用の形式で記述されるため、比較的容易に抽出できます。そこで、このような文書における構造的特徴を考慮し、文章中に単独で登場する変数記号を自動で抽出するアルゴリズムを作成しました [3]。

一方で、変数記号や数式とは異なり、変数定義を文中から抽出するには文の意味を扱う必要があります。自然言語処理の分野では、記号に対する定義を抽出する方法として、指定したパターンに合致するものを抽出するルールベース手法 [4]、自作の特徴量と機械学習を用いる手法 [5, 6]、ニューラルネットワークモデルを用いる手法 [7–9]などが提案されています。複数の先行研究 [8, 9]において、BERT (bidirectional encoder representations from transformers) [10]というニューラルネットワークモデルを用いる方法が高い性能を達成していたことから、私たちも、BERTを用いた変数定義抽出手法の開発に取り組みました [11, 12]。

Fig. 2は、提案した変数定義抽出手法の概要図です。この方法では、BERTと同様の構造を有するモデル（以降、BERTモデルと呼ぶ）に文を入力し、文中の各トークン（単語より短い文の構成要素）が変数定義の開始位置と終了位置である確

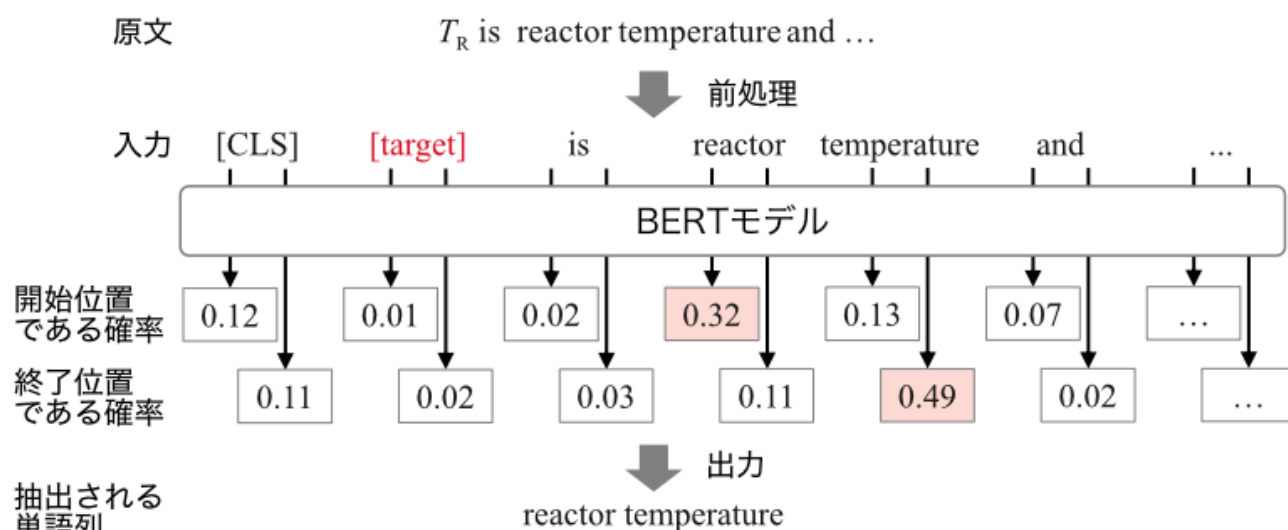


Fig. 2 BERT モデルを用いた変数定義抽出手法 [13]

率を計算します。そして、この2つの確率の和が最大となるトークン列を変数定義として出力します。

本手法で使用するBERTモデルのパラメータを決めるには、変数記号と対応する変数定義を含む文を複数用意する必要があります。そこで、製造プロセスに関連する論文45報に対して、変数記号自動抽出アルゴリズムを適用し、抽出された変数記号に対応する変数定義を手動で抽出することで、変数記号と変数定義の位置がラベル付けされたデータセットを作成しました [3]。このデータセットを用いてDeBERTaV3_{large} [14]というBERTモデルを学習（パラメータを決定）したところ、正解率85.4%を達成しました。これは、他分野における変数定義抽出 [8, 9]と比較しても高い性能です。ただし、AutoPMoBの実現に十分な性能かどうかは未確認であり、情報抽出以外の工程と組み合わせることでのどのような問題が起こるのかを確認する必要があります。

4. 抽出した情報の表記統一

複数の文献から抽出した情報が同一の物理現象を表現しているかを判定することも、AutoPMoBにおいて重要な技術要素です。異なる文献において、同じ物理量・現象が異なる変数・数式で表現されることがあります。例えば、 $dC/dt = k_1 C$ と $\dot{C}_A = k_0 C_A$ は、同一の現象（一次反応）を表現しています。このような文献による表記の違いに対処するため、変数と数式を順に表記統一する手法を提案しまし

た。

変数については、変数記号と変数定義のペアをBERTモデルに入力して、それらが同じ意味（同義）かどうか（同義であれば1、そうでなければ0）を出力する手法を提案しました [15, 16]。この手法で5種類の製造プロセスに関する論文に登場する変数ペアの同義性を判定したところ、90%程度の同義ペアを正しく判定することができました。しかし、製造プロセス単位で性能を確認したところ、同義ペアをすべて非同義であると判定したプロセスも確認されました。高い性能を達成することが難しい原因の一つとして、同義ペア（正例）の数が非同義ペア（負例）の数に対して極端に少ないことが挙げられます。各製造プロセスに対して、十分な数のサンプルを用意することは容易ではないため、似たデータを使用したり学習方法を工夫したりすることで性能向上させることが今後の課題です。

数式の同義性判定については、数式が従う表記規則は自然言語とは異なるため、BERTモデルではなく計算機代数システム（computer algebra system; CAS）を用います。CASの代表的ソフトウェアであるMathematicaやMapleでは、Simplifyのように数式を単純化して比較する関数が定義されており、2つの数式の同義性を判定することができます。しかし、複数の数式（数式群と呼ぶ）同士の同義性判定には対応していません。そこで、物理モデルの記述によく用いられる微分代数方程式で構成される数式群を対象に、CASを用いて代数的な同義性を判定する手法を開発しました [17]。この手法では、まず、変数消去によって2つの数式群に登場する変数の種類を統一します。続いて、片方の数式群に存在するすべての数式に対して、同義の数式がもう一方の数式に存在するかを確認します。すべての数式に対して同義の数式が存在すれば、2つの数式群は同義であり、そうでなければ非同義であると判定します。プロセス制御に関する書籍に登場する物理モデルを参考に50組の物理モデルのペアを作成し、それらに対して提案手法を適用したところ、すべてに対して同義性を正しく判定できました。

このように、抽出した情報の種類に応じて使用するツールを使い分けることで、正確に同義性を判定して表記を統一する技術を実現することを目指しています。ただし、いずれの方法も実運用に十分な性能に達しているかは未検討です。情報抽出などの工程と組み合わせて、実応用を想定したケーススタディを通して検証

と改良をすることが今後の課題です。

5. おわりに

本稿では、物理モデルを自動で構築するシステム（AutoPMoB）の全体像と、その構成要素である情報抽出と表記統一の手法を紹介しました。情報抽出では高い性能を達成していますが、AutoPMoBを実現するには、各手法の性能向上と、それらを統合する仕組みの整備が必要です。各手法の詳細や他の手法の説明に関しては、本文中で引用している論文や解説記事 [13]を参照ください。

最近では、文書を対象とする研究において、ChatGPTやClaudeといった大規模言語モデル（large language model; LLM）を用いたサービスが有力な選択肢となります。LLMは自然言語処理の分野で急速に進化しており、文書からの情報抽出や数式の理解においても高い性能を発揮しているためです。現在、LLMを活用したアプローチも検討しており、より効率的な物理モデル構築システムの実現を目指しています。

「複数の論文や書籍から物理モデルを作るAIを実現することはできるのか？」という問いに対し、研究を始めた当初は、難しく、実現には数十年を要すると考えていました（[18]では、所属研究室の加納教授が「妄想全開」と表現されています）。しかし現在では、「数年以内には実現できるはず」と答えられる段階に近づいています。今後も、急速に進展するAI技術を活用しながら、AutoPMoBの実現に向けて研究を推進していきます。

本研究はプロセス産業を主な対象としていますが、機械系を含む他の工学分野にも有用であると考えています。本稿に関心を持たれた方は、ぜひご意見やご要望をお寄せいただければ幸いです。

最後に、自由に研究を進める環境を提供してくださっている加納教授および加納研究室の皆様、そして共同研究者の皆様に深く感謝申し上げます。

参考文献

[1] <https://www.scej.org/general/whatisce.html>

[2] S. Kato, S. Kim, M. Kano, T. Fujiwara, and M. Mizuta: Gray-box modeling of 300 mm diameter Czochralski single-crystal Si production process; J. Cryst. Growth, Vol. 553, 125929 (2021)

- [3] S. Kato and M. Kano: VARAT: variable annotation tool for documents on manufacturing processes; J. Chem. Eng. Jpn., Vol. 58, 2454461 (2025)
- [4] R. Pagel and M. Schubots: Mathematical language processing project; CEUR Workshop Proceedings, Vol. 1186 (2014)
- [5] J. Lin, X. Wang, Z. Wang, D. Beyette, and J.-C. Liu: Prediction of mathematical expression declarations based on spatial, semantic, and syntactic analysis; Proceedings of DocEng'19 (2019)
- [6] M. Schubotz, L. Krämer, N. Meuschke, F. Hamborg, and B. Gipp: Evaluating and improving the extraction of mathematical identifier definitions; Proceedings of CLEF 2017 (2017)
- [7] Y. Stathopoulos, S. Baker, M. Rei, and S. Teufel: Variable typing: Assigning meaning to variables in mathematical text; Proceedings of NAACL-HLT 2018, pp. 303–312 (2018)
- [8] S. M. Lee and S. H. Na: JBNU-CCLab at SemEval-2022 task 12: machine reading comprehension and span pair classification for linking mathematical symbols to their descriptions; Proceedings of SemEval-2022, pp. 1679–1686 (2022)
- [9] N. Popovic, W. Laurito, and M. Färber: AIFB-WebScience at SemEval-2022 task 12: relation extraction first - using relation extraction to identify entities; Proceedings of SemEval-2022, pp. 1687–1694 (2022)
- [10] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova: BERT: pretraining of deep bidirectional transformers for language understanding; Proceedings of NAACL-HLT 2019, pp. 4171–4186 (2019)
- [11] S. Kato, M. Yamamoto, K. Nagayama, and M. Kano: Two-stage fine-tuning for variable definition extraction from chemical process-related papers; DAFUSAI (2024)
- [12] K. Nagayama, S. Kato, and M. Kano: Data augmentation method utilizing template sentences for variable definition extraction; Proceedings of NLDB 2024 (2024)
- [13] 加藤: プロセス産業における自然言語処理を用いた物理モデル自動構築; システム/制御/情報, Vol. 69, pp. 184–189 (2025)

- [14] P. He, J. Gao, and W. Chen: DeBERTaV3: improving DeBERTa using ELECTRA-style pre-training with gradient-disentangled embedding sharing; Proceedings of ICLR 2023 (2023)
- [15] S. Kato, K. Kanegami, and M. Kano: ProcessBERT: a pre-trained language model for judging equivalence of variable definitions in process models; IFAC-PapersOnLine, Vol. 55, pp. 957–962 (2022)
- [16] 張, 加藤, 加納: ProcessBERTの性能改善に向けた数学的表現の構造的情報のトークン化手法; 言語処理学会第29回年次大会発表論文集, pp. 2940–2944 (2023)
- [17] S. Kato, C. Zhang and M. Kano: Simple algorithm for judging equivalence of differential-algebraic equation systems; Sci. Rep., Vol. 13, 11534 (2023)
- [18] https://note.com/dr_kano/n/n443fe849f8d9

Profile

京都大学大学院情報学研究科情報学専攻人間機械共生系講座ヒューマンシステム論分野 助教

1995年、岐阜県羽島郡に生まれる。2017年、京都大学工学部工業化学科を卒業。2019年に同大学大学院工学研究科化学工学専攻修士課程修了。2022年、同大学院情報学研究科システム科学専攻博士後期課程修了後、博士号取得。2022年から京都大学大学院情報学研究科助教。

研究室HP : <http://human.sys.i.kyoto-u.ac.jp>

第66回前期京岬会（昭和33年卒業同期会）報告

中村弥寿家（S33/1958卒）

本年度懇親会を5月28日に彦根市の料亭で7名の参加をえて実施した。

参加者は89～90歳の高齢ながら意気だけは健勝で長時間の交友を楽しんだ。

会の近況報告では本年は従来から出席の2名が鬼籍に入りました。自動車運転は3名が継続中。

各人の近況報告とフリーデスカッションを行った。



写真後列	（左より）	倉田	岸本	中村	小澤
写真前列		造田	梅本	新田	

近況報告

中村：季節の花めぐり観光が好きで今年は桜見物（前橋市、高崎市）つつじ見物（館林市）をしました。近場の公園で多くの花見物。

小澤：大阪倶楽部の講演会（毎週開催）に参加、趣味の書道、海釣り、料理、U-Tubeでカラオケ等を楽しむ。

新田：高等数学の研究継続中。

倉田：自由詩創作、近々詩集を発刊予定。

岸本：楽器（フルート）を演奏、小楽団で演奏会も楽しんでいる。

造田：喫煙継続中、小旅行を楽しんでいる。

梅本：原因不明の体調不全に悩まされており、ストレス緩和のため温泉旅行をしている。

自由討議

- ・ 我国農業政策の課題、問題点
- ・ トランプ外交政策への異見
- ・ 我国工業力の評価、期待自動車その他、機械力（高速電車、動力設備）半導体、データバンク等

- ・ 日本製鐵のUSSのM&Aの今後の行くえ。黄金株の評価等

翌日のゴルフ参加者は小澤、岸本、中村の3名となり今回以降プライベートの行事となった。

本懇親会もゴルフが無くなり総会とし5月、10月の年2回の総会をすることになった。