

研究の手引き (平成 25 年度版)

神奈川大学理学部情報科学科 田中 賢

概要

新たに研究室に入ってくる3年生とこれから卒研に取り組もうとする4年生、院生を対象に、研究への取り組み方をまとめた。まず、研究と勉強の違いをはっきりさせよう。研究活動では自分の思考力と判断力を頼りに、「自分自身のエンジンで」走れなくてはならない。走るための原動力は、疑問を抱く力である。

実際に研究テーマに取り組む際の思考のヒント、文献の調べ方、ゼミへの参加、プレゼンなど、HowTo的な内容もまとめてみた。この中には、社会に出て企業で働く際にも必須の事項が含まれているが、社会に出たら教えてもらえる機会はおそらくないだろう。

1 研究とは

1.1 勉強と研究の違いについて

勉強と研究の違いは、答えや問題があらかじめ与えられているかどうかにある。研究は、未知の対象を探り、知り、あわよくば利用することを目指す行為だから、問題や答えがないのは当然のことだ。小学校以来、学生諸君が求められてきたことは、あらかじめ用意された講義内容を正確に早く身につけて行くことだけだっただろう。試験においても、「決められた範囲のものをいかに正確に効率的に身につけることができたか」をもって常に評価されてきたわけだが、そういう評価尺度はそもそもこの世界では捨てなければいけない。時間がかかっても本当に大切なことを見出せることこそが唯一重視されるのだ。

1.2 研究に取り組む姿勢

研究室で研究をする上で大切なことは、研究に関わる全てのひとが、自分の考えを自由に述べることができることだ。教師と学生という関係を意識するのは余計なことで、その問題について一番正しい確かな意見だけが尊重されなければ研究は成り立たない。議論をする際も、学生さんは、自分のいってることが不完全であるとか、間違っているんじゃないか、ということに気にする必要はまったくない。そんなことは研究をしていれば日常茶飯事で、僕が言ってることも間違いだらけなのが普通なのである。研究室では、皆が気軽に自分の意見を言えるような雰囲気を作ることが何より大切だということ、学生同士も僕も常に心がけたい。

このことは言い換えれば、僕に聞いても正解は教えてもらえない、ということだ。だって、僕も正解を知らないのだから。だからといって、僕に何かを期待したり求めたりしてはいけないということではなくて、困ったとき、行き詰まったときは、僕のところに気軽に足を運んで、一緒に悩むのがよい。一緒に悩むことの効果はしばしば期待以上のものがあって、言葉を交わしながら一緒に悩むことで誰かの頭の中からよい考えが浮かぶ、ということは過去に何度も経験してきたことだ。研究もコミュニケーションが大切なのである。

これは研究に限らないが、何か話をする際「メモをとること」は非常に重要である。ほとんどの人は天才ではないので、話した内容を片っ端から忘れていく。大切な研究のヒントや、指示された事柄がいつのまにか忘れ去られて、「そういえば、あのことってどうなった?」「あ...」となりがちである。社会に出たら、それは重大な信用問題につながることもある。メモをとるのは、「それを後から確かめるため」というよりは「メモをとることでその場で記憶にとどめる」のが目的である。やってみると分かるが、一度メモったことは大概忘れないものだ。例年、最初のゼミのときに「メモをとることは大切だよ。」という話をするのだが、その話をしているときにメモをとる学生は一人もいないのだ。たった今からメモをとる習慣をつけるべし。いつやるの、いまでしょ、である。

2 研究テーマへの取り組み方

研究テーマは、例年僕がいくつかの候補を出して、各自の関心に応じて学生同士で相談して選択する方法を取っている。でも、中には自主的にやりたいことがあるという学生もいるだろうし、自分で研究テーマを探したいという学生もいるだろう。田中研では、そういうことを歓迎する。研究とはそもそも自分の興味を探究することの延長で、誰かに与えられてやるものではない。もし自分でやりたいことがあったら、できるだけ本人の希望に沿うよう配慮したい。自分で研究テーマを作ってみたい、という学生は、この文書の付録の「研究テーマの作り方」を参照されたい。

2.1 研究のヒント

研究テーマが決まったら、設定された問いの解を目指して、一步一步進んでいかなければならない。研究テーマによって突き当たる壁は様々だが、ここでは研究を進める際に知っておくとよいいくつかのヒントを述べる。

1. 例を用いて考える

アルゴリズムを考えるには、まず個々の事例においてどのような結果が望ましいかかを考えるのが先決だ。例えば、フィルタリングルール最適化のアルゴリズムを考えるときも、いままでうまく最適化できなかったルール集合に対して、こんな風になったら都合がよい、という例が必ずあるはずだ。問題の特徴を最もよく表す例を見つけて、それについてどのように結果を得ればよいか考えてみよう。よい例は、解法のポイントを捉えるのに役立つので、他人に研究内容を説明するのに後々まで役立つ。

2. サイズを小さくして考える。

対象を単純化すると本質が見えることがある。視覚化が可能になり、理解も容易になる。

例)

ルールが2つだけのルール集合を考える。そこから3つ、4つと増やしていったら考えてみる。

3. 問題の前提をゆるめて考える。

自分にとって都合のよい特定の条件を前提にして問題を考えてみる。あるいは、都合の悪い条件を意図的に外してみる。解が見出せる所まで前提を変更していく。解が得られたら、そこから逆に一般の状況を考えていく。

例)

デフォルトルール以外、全てが許可ルールだったらどうなるだろう。あるいは、全てのルールが独立だったらどうなるだろう、など。

4. 特殊な場合・極端な場合について考える。例を用いて考える場合の一種だが、これはむしろバグ出しに有用なことが多い。

人間はあまり記憶力がよくないものだから、よいことをおもいついてもどんどん消えさって行く。自分が何かを覚えていたことすら忘れてしまうのだから始末が悪い。研究ノートを作って、思い付いたこと、疑問に思ったことが消え去るのを防ごう。

2.2 研究に行き詰まったら

適当な問題が見つかって、すんなりと答えが出てくるのは最初の段階だけで、壁に突き当たるのがよくある。安易に行うことを進めるわけではないが、「問題解決の糸口」が見つからないなら、問題を別の方向に発展させてみるのは1つの手である。問題そのものをかえてしまうということだ。

例)

遅延を最小にするアクセスリストを求めることは簡単ではない。それなら、まずルール数を最小にするにはどうしたらよいか、について考えてみる(ルールが少なくなれば普通は遅延も小さくなるものだから)。

研究というのは自由気ままに進めてよいものだ。最初にも述べたように、疑問に思ったこと、気になった点、困った点、その他話したいことがあれば、僕を気軽につかまえてほしい。

3 文献調査

研究の価値は、新規性、有効性、信頼性の3点で計られる¹。既知の事柄は例えそれが有効であっても、信頼できても価値はないわけだから、新規性があることが全ての研究テーマの必須条件になる。

学生諸君は自分が取り組もうとしているテーマについて考え始める前に、その分野で何が既知か」をきちんと調べる必要がある。せっかく研究結果を出しても、それが既知のことであつたら価値はゼロと見なされてしまうからだ。情報を集めることは、そのままその研究テーマについて深く知ることにつながり、自分がテーマに取り組む際の有用な手がかりともなる。

情報を集める代表的な手段をあげてみる。

1. 論文誌

論文誌にはいくつかの種類がある。

- 学会論文誌 (Transaction) 学会が学会員の研究発表のために発行する論文誌
- 専門雑誌 (Journal) 出版社が専門家のために発行する雑誌
- 研究報告書 (Report) 研究会の予稿集
- 国際会議議事録 (Proceedings) 国際会議の予稿集
- 大会予稿集 (Proceedings) 全国大会・地方大会の予稿集

学会論文誌、専門雑誌は、複数の専門家の査読を経て掲載が許可されるので、内容の信頼性が高い。研究報告書と大会予稿集は、査読がない分信頼性が低く玉石混淆だが、掲載までの時間が短いので最新の情報を得やすい。国際会議議事録はその中間である。

関連する全ての論文誌について、過去10年分程度の文献を title と abstract を中心に丹念にあたってみよう。関連のありそうなものは、コピーをとってファイルする。このとき必ず出典を書き留める。あとで論文を書くとき、参考文献リストで必ず必要になる。田中研に關係する論文誌をいくつか上げておく。

インターネット

- 電子情報通信学会論文誌 (B 分野, D 分野)
- 上記の英文誌
- 電子情報通信学会技術研究報告書 (IA インターネットアーキテクチャ)
- 情報処理学会論文誌
- *IEEE Transactions on Communications*
- *ELSEVIER Computer Communication*
- ACM(アメリカコンピュータ学会)の各種ジャーナル, トランザクション

計算・言語・学習

- *Information and Computation*
- *Information and Control*
- *Theoretical Computer Science*
- Discrete Applied Mathematics
- 人口知能学会誌

ニューラルネットワーク

- 電子情報通信学会論文誌 (DII 分野)
- 電子情報通信学会技術研究報告書 (NC ニューロコンピューティング)
- 情報処理学会論文誌
- 計測自動制御学会論文誌
- *Neural Networks*
- *Neural Computation*
- *IEEE Transactions on Neural Networks*
- *Proceedings of IEEE International Joint Conference on Neural Networks(IJCAI)*

2. インターネット

インターネットで情報を得るには、WWWの検索エンジンが最も手軽だ。論文の著作権が学会にあるため、通常 Web で手に入るのは文献の概略までだ。必要な文献は図書館から取り寄せる。主要な研究分野には、大抵メーリングリストがあるので参加すると便利だ。

¹ 3つの要素を十分に備えた論文を原著論文という。

3. 文献データベース

大学の図書館では文献データベースを利用できる。1年程前の情報までが収録されている。キーワードで瞬時に検索できるため重宝するが、取り出せる結果がキーワードの選定に大きく依存するため、注意する必要がある。1つの単語が、複数形と単数形で別々に分類されている場合もあるし、同じような意味の別の単語で分類される場合もある。事前に考えられるキーワードのリストを作っておくのがよい。検索は無料だが複写するには費用がかかる。院生は研究費で払うことができるが学部生は払えないので院生と一緒にいって名前を借りて複写すること。

4. 専門書

情報としては古いが、代表的な研究結果が簡潔に収録されていて勉強するにはよい。ここの図書館には関連書籍がほとんどないので、図書館の検索システム OPAC で探そう。ただ、研究に使う本は常時手元にあった方がよいので、内容を確認したら購入して研究室の本棚に置いた方がよいかもしれない。研究室の本棚には新しい参考書も随時補充しているが、必要な本がある場合は小生に確認の上書籍部で購入申し込みをすればよい。専門書については、付録で再度触れる。

コンピュータ関係の本の品揃えが充実しているのは、藤沢のジュンク堂だろう。新宿のジュンク堂の方が品揃えはよいと思う。なお、日本で一番大きな本屋は池袋のジュンク堂らしい。領収書があれば、立て替えて払ってあとから大学にお金を請求することもできる。新宿の紀伊國屋書店は、大学の書籍部との関係で、代金後払いで持ち帰ることができる。1年に1回、学生と一緒に本の買い出しに行く。研究に関連する本ならなんでも買ってよい。

参考になりそうな論文が手に入ったら、何度も読んで完璧に理解しよう。1つの論文を理解するためには、そこで引用されている他の論文を合わせて読む必要がある場合も多い。必要な論文は再度取り寄せるようにする。

集まった文献は、BiBTeX でリストを作っておく。入手した論文は、内容ごとにファイルして表紙をつける。以後、新しい文献を手にいれたら、そのつどこれに追加していく。これも、あとあと論文を書くときに役に立つ。論文の読み方については、付録で再度触れる。

4 ゼミについて

2, 3週に1回、研究の進捗状況を皆の前で発表する。毎回毎回、研究の成果を求めているのではなく、前回から今日までにどんなことを考えてきたか、やってきたか、そのことの報告をする。

4.1 ゼミ資料の書き方

ゼミ資料の文章や図表は、そのまま卒論の一部になることが多い。日頃からきちんとしたゼミ資料を作るよう心がけていないと、最後に困ることになる。

ここで述べることは、以後のプレゼン資料の作成や卒論の執筆の際にもそのままあてはまる。ゼミ資料といえども、他人に自分の行ったことを理解してもらうために書くれっきとした論文の一種だからだ。この「他人に理解してもらう」ということを、まずきちんと肝に銘じておく必要がある。例えば、ゼミ資料の中で

聞き手の知らない用語や概念を用いて説明してはいけない

聞いている人は、知らない用語が出てきた段階で、それ以後の話は全てわからなくなってしまう。聞き手が知らない用語を用いるときには、必ずあらかじめ用語の意味を説明しなくてはならない。聞き手、読み手が、一本道をたどるように内容を理解できる文章にしなくてはならないわけだ。

効率的に内容を理解してもらおうとすると、ゼミ資料の構成も自然と1つの形に納まってくる。それは以下のような構成だ。

1. 題名
2. 概略
3. 準備
4. 前回までの経緯・問題点
5. 本論
6. まとめと今後の課題

7. 参考文献リスト

8. チェックリスト

題名の決め方

題名は、その論文を手にとった人が研究の概略を人目で把握できるようなものにすべきだ。例えば、こんな題名は駄目である。

“フィルタリングルール最適化問題に関する基礎的研究”

これでは、基礎的と言っているだけで範囲が広すぎてよくわからない。同じ理由で「高速」とか「高精度」とかいう漠然とした形容詞も題名には用いるべきではない。そういう題名をつけると、中身がない研究だと思われるのだ。

“従属ルールを含むフィルタリングルール最適化問題に関する基礎的研究”

このように書くと、大部わかってくるが、さらに具体的に書きたい。

“従属ルールを含むフィルタリングルール最適化問題の多項式時間アルゴリズムについて”

これくらい書けば、読み手は「ああ、なるほど」と思ってくれるわけだ。

概略

ゼミ資料の内容を数行で書く。ここでは細かな内容は書かず、自分の研究テーマの中で「どんな疑問について」「どんな所に着眼して」「どんな検討をし」「どんな結論を得たか」だけを簡潔に書く。

準備

ゼミの参加者をあらかじめ予想して、まず研究テーマを理解するために必要な予備知識の説明をする。新4年生がいるなら、一番基本的な用語や概念から説明しなくてはいけないだろう。新人でなくても、参加者は発表者とは違って「普段はその研究テーマのことは考えていない人達」だから、2週間もすれば粗方忘れてしまう。毎回、簡潔に予備知識を復習するようにすべきだ。予備知識を与えた所で、自分が取り組んでいる研究テーマについて概説する。

前回までの経緯・問題点

研究テーマについて、前回までにどんな問題に取り組み、どんな結論を得、何が問題になっていたかを概説する。

本論

前回から今回までに自分が考えたことを書く。

まとめと今後の課題

今回明らかになったこと、次回以降に検討すべき課題を書く。

参考文献リスト

ゼミ資料の中で引用した文献を全てリストアップする。文献リストには、決まった書き方があるのでそれに従う。

チェックリスト

何度もゼミを続けて行くと、何が論点だったか段々わからなくなって「研究がだらけてくる」。自分で気づいた点や、ゼミの参加者に指摘された問題点をチェックリストに追記していき、それがどこまで解決されたかを × 式で管理する。

ゼミ資料の形式は以上だが、問題はこれらの項目を「読み手が理解できるように日本語としてきちんと表現できるか」である。これについては、この分野で定番と言われる指南書がいろいろあるので、それをよく読んで欲しい [3]。以下では、最低限のことについてだけ触れよう。

4.2 正しい日本語を

”フィルタリングルール最適化問題における最も最大の問題点は、Lの最小値が最大とならないわけではない。”

… いるんです。こういう人が。日本人なのに日本語が使えない人が。こういう似非日本語が来ると、かなり精神的にやられます。私がやられないために、学生諸君はまず「理科系の作文技術」木下是雄（中公新書）を必読すること²。

それはともかく、日本語文として理解できるために、まず以下の項目に注意しよう。

1. 主語と述語の対応がとれているように。主語が2つ以上あったり、文の途中で主語が入れ替わってしまうような文は避ける。

2. 指示代名詞がわかる、つまり「この」と言ったら、「どの」かが分かるように

以下は論文として避けるべき表現です。

1. 「… れる」「… られる」「… される」といった受動的表現

2. 曖昧な表現

例) 「ほとんど不可能」、「… と思われる」、「比較的大きな効果」

3. 括弧を用いた説明

例) 結合重みの値 (正か負か, 大小など) によって …

日本語には受動的表現が多いが、動作の主体がわからないので説明として不十分になってしまうことが多い。要するに、曖昧なことは言ってはいけないのだ。

名前をつけよう

例えば、以下のような文は読むに耐えるだろうか。

”11 を評価するルールを通過したパケットが 1* を評価する出力するルールに到達すると、11 を評価するルールと 1* を評価するルールによって無駄に評価されるため、1* を評価するルールをあらかじめ 11 を評価するルールと併合し … ”

このようなときは、登場する要素にあらかじめ名前をつけるとすっきりする。

”11 を評価するルールを R_1 、* を評価するルールを R_2 とする。 R_1 を通過したパケットが R_2 に到達すると R_1 と R_2 によって無駄に評価されるため … ”

図を使おう

上述のような文を理解しようとするとき、人は頭の中で一生懸命絵を描いている。 R_1 や R_2 を言葉で説明するだけでなく、図を添えると一層わかりやすくなる。

表を使おう

以下のような文はどうだろうか。

”実験 1 の場合、サイズ 20 のネットワークの成功率は 0.10、サイズ 10 の場合の成功率は 0.23 となった。実験 2 においては、サイズ 25 のネットワークを加え、同様の実験を行い … ”

それぞれの実験やネットワークのサイズについて、成功率を「比較」したいのなら、横軸に実験、縦軸にネットワークのサイズをとった表を添えるべきだ。

なお、句読点・句点は、電子情報通信学会の基準に従い「、」と「。」を用いる。

4.3 ゼミの中で心がけること

ゼミは頭を働かせて疑問を投げ掛け合う場なので、出来るだけリラックスしたざっくばらんな雰囲気の中で行うのがよいと思っている。参加者はどんどん質問をしよう。どんな小さなことでも簡単そうなことでも構わない。自分が知りたいと思っていることは、他の人も知りたいと思っている可能性が高いのだ³。どんなことでも知れば面白くなるし面白くなれば疑問やアイデアが湧いてくるものだ。

²この本はとても有名な本で、これを読んでいない技術者は潜りといってもいいくらいです。

³講義でもゼミでも、学生から出た質問について、「なーんだ、そんな簡単なことも分からないのか。」と感じたことは今だから一度もないです。

最初にも書いたが、発表者はとにかくこまめにメモを取ろう。覚えているつもりでも、人はどんどん忘れていってしまうものだ。悲しいことに、人は「自分がメモを取ったということすら」忘れてしまう。最初はとにかく全部書き留める位でちょうど良い。メモの効能には「書いて覚える」ということもあるからだ。議論の中で重要だと思った点や、方針らしきものが出たときには、 で囲ったり下線を引いて、次回チェックリストに忘れずに追記しよう。

5 プレゼンテーション

プレゼンテーションについてもよい指南書がたくさんあるので、ここでは要点を述べるにとどめる。

5.1 プレゼンテーションシートの作り方

卒研発表会でよく見かける光景だが、最悪のプレゼンは、シートが映し出されたスクリーンを発表者が棒読みする、というものだ。プレゼンシートは、聴衆が発表の流れを把握し、内容を容易に理解するための「聴衆のための補助的な情報」で、発表者のメモではないことに注意して欲しい。

プレゼンテーションシートの構成は以下ようになる。

1. 表紙
研究テーマ名、所属と研究室名、名前を書く。
2. 研究の背景
3. 研究の目的
4. …
5. まとめと今後の課題

各シートの始めにはそのシートで説明する内容のタイトルをできるだけ1単語で、無理なら「 」程度の長さでつける。各シートの内容は10行程度が上限なので、説明時間は1枚1分が大まかな目安になる。15分の発表なら、表紙を0分として15+1枚位にまとめるようにする。

余程専門家が集まっている場でない限り、聴衆は学生諸君の研究内容を聴くのははじめてだ。複雑な内容を詳細に説明されても実際にはほとんど理解できない。研究内容を全て理解してもらう事を目指すのではなく、「どんな問題を扱うのか」「なぜそうしたのか」といった、全体的な流れのポイントを抑えていくようにするのが大切だ。これについては、次節の発表の仕方でも再度述べる。

シートの内容について、発表者は責任を持たなくては行けない。聴衆から質問が出たとき、知識が不確かできちんと説明できないようなことは最初から書いてはいけない。自分の言葉で完璧に説明できる事柄だけをのせるようにする。

シートを作る上で特に大切な点は、「聴衆に読ませる内容は最小限にせよ」、ということだ。聴衆は、発表者の話を聞きながら同時にシートも読まなくてはならない。読みながら聞くということは普通の人間には不可能で、実際は読むのと聞くのを繰り返しているのだ。読むことにあまり時間をとられると、聞く方まで手がまわらないことになり、内容の把握が疎かになる。結局、よくわからない発表だという印象だけが残ってしまう。同じ内容を書き表すのでも、極力簡潔な書き方をしよう。個別の注意点は以下である。

- 内容を理解する上でポイントになる概念は、最初のうちに例を交えてじっくり説明する。
- できるだけ図を挿入する。図の説明は口頭で行い、シートの中での説明は最小限にする。
- 式の導出を示すなら、導出の手順のポイントを言葉で記述する。
- 論理の流れ「…ゆえに…」は、 記号で。
- シートの中では、文は書かず全て体言止めにする。
- 何が可能になったのか、何がわかったのか、最後に結論をはっきり書く。
- まとめと今後の課題は箇条書に。
- 背景と目的は、発表時間に応じて1枚にまとめてもよい。
- 同様に、まとめと今後の課題は、別シートにしてもよい。

6 発表の仕方

プレゼンテーションは、自分の頭の中から、そのつど言葉をつむぎ出しながら、伝えたいことを分かりやすく聴衆に理解させる行為だ。用意した文章を読み上げる行為はプレゼンテーションではない。初めての発表のときは、原稿を作って全体を見通してみるのもよいが、そういうものではないということをまず心にとめておきたい。

分かりやすい説明とはどんな説明だろう。それは、話の流れが分かる説明だ。「こういうことが問題になっています」「従来は、こういう方法が用いられてきました。」「ところが、そこではこういうことが問題になります。」「そこで、それを解決するために、こういうことを考えました。」「この方法だと、これくらいうまくいきます。」「ただし、こういう問題が新たに起きてきます。」「今後は、この問題を解決するのが課題です。」という具合に、接続詞を使って各々のシートを繋いで1つの物語ができるように説明をしていくわけだ。

6.1 発表時の注意

発表時の注意点は以下である。

- 説明は、研究の背景からいきなり入ってよい。通常、題名や所属、氏名は座長がすでに紹介しているからだ。紹介がなかった場合には、それらから始めることになる。
- シートを読まないことはもちろんだが、「研究の背景として…」「研究の目的として…」「まとめとして…」という言い方で始める発表は、結局読んでいるだけという雑な印象を与える。研究の話をするのだから「研究の」はいらないだろう。背景や目的の話をしていることはシートを見れば分かる。「として」はそもそも不要。語り始めてからシートに目をやるようにすれば、このようなことは起きない。「正則言語問題は、未知の正則言語の正例と負例が与えられたとき、それに合致する状態数最小の有限オートマトンを求める問題です。」「今回の研究でということがわかりました。」「今後は について考えていきます。」という具合に、キャスターがニュースを語るような、正確で自然な日本語で説明しよう。
- 自分の研究を説明するのだから、主語は基本的に「私は」である。当然言い回しは、「(私は) を と置きます」「(私は) 次に をします。」となる。受け身形の言い回し、自動詞(～になります)はできるだけ使わない。
- 言葉の埋め草(～ための、～のような、～がわかります、など)を入れず、極力言葉を少なくすると聴きやすくなる。
例)「おさえるための方法」「ための」はいらない。「おさえる方法」でよし。
例)「最小化することを試みます」「ことを試みます」はいらない。
「最小化します」でよし。
例)「改善されることがわかります」「ことがわかります」はいらない。
「改善されています」でよし。
- 式の説明はクドクド言わず最小限にする。もし説明が必要なら、「 に をかけたものを…」など、式の構成を順をおって指し示しながら説明する。
- 定義の表現法は「 を表すために、 を定義します。」とする。 は意味を表す直観的な概念で、 はそれに与える形式である。
例)「ネットワーク機器の負荷を表すために、関数 $L(R)$ を以下のように定義します。

6.2 質疑応答

発表後の質疑応答で「はまる」のは、だいたい以下のようなパターンだ。

- 何を聞かれているのかわからなくて答えそのものが的外れになるケース。聞かれている内容が分からないなら、逆に「質問の主旨はこのようなことですか？」と何度でも問い返そう。そうすることは何ら問題ない。
- 答えが問いと対応していなくて食い違ってしまう、対応する答えを引き出そうとして、やりとりが繰り返されてしまうケース。例えば、「～はできますか？」という問いに対する答えは、「できます。」「できません。」のいずれかである。「～はいくつですか？」に対する答えは、「(数字)です。」しかありえない。「～はどんな点ですか？」に対しては「～という点です。」が答えだ。要するに聞かれたことに答えましょう、ということだ。答えに但し書きをつけることはかまわないが、その場合はそれに当てはまらない場合についての答えも必要だ。「～であればできます。

それ以外の場合については、現時点では不明です。」など。全くわからないならば、「わかりません。」というのも立派な答えである。

- きちんと理解できていない概念や用語について質問を受け、答えられない。
例)「ブール半環の … 」「ブール半環って何ですか?」「 … 」「じゃ環って何ですか?」「 … 」
こういう
こういう場合はもうどうしようもない。「すみません、勉強不足でしたので、今後不明点を無くすよう努めます」とでもいう。こうならないように、プレゼンの内容は「全て自分がきちんと理解していること」だけに限定しなくては行けない。分からないことを無闇に並べてもいいことは何もない。
- 話が噛み合わなくて収束の目途が立たない。そういう場合は、「時間が過ぎているので、出来ればあとで個別に議論させてください。」と提案する。発表が終わったあと相手の所に行きつくり話をする。
- 提案手法や問題設定の根幹にかかわるような質問で、やりとりしてもうまく やり返せない。そういうときは「今はうまくお答えできませんが、大切な指摘だと思いますので、今後よく検討したいと思います。貴重なご意見をありがとうございました。」と収束させる。

7 論文(修士論文, 卒業論文, 卒研要旨)の書き方

論文を読む可能性が最も高いのは、実は諸君の後に新たに入ってくる4年生だ。諸君がやり残した課題にさらに取り組んだり、発展的な課題に取り組むための最初の材料が卒業論文である。だから、卒業論文は、その分野について全く知らない人が読んでもきちんと理解できるよう、研究の背景から参考文献に至るまで、詳細に分かりやすく記述されるべきだ。

卒業論文に限らず、論文は以下のような構成で記述する。もちろん、これらがそのまま章題になるわけではないので注意して欲しい。

1. 題名と著者の所属・氏名
2. 概略
3. はじめに
4. 準備
5. 本論
6. 議論
7. まとめ
8. 参考文献リスト

ゼミ資料の書き方に準ずる部分も多いので、ここでは論文としてまとめる際の注意点をのべる。

題名

題名は、卒研発表会の発表題名と同一である。所属は 大学 学科 研究室とする。

概略

論文を読む人は、まず「題名」を読む。そして、その題名に興味を感じたら「概略」を読む。さらに興味を感じたら「まとめ」を読み、そこではじめて本論を読むべきかどうかを判断する。だから、題名や概略の役割は重要で、論文を手にとった人が、研究の概略を速やかに把握できるようにしなくては行けない。概略がダメだと、その論文は誰にも読んでもらえないことになる。

「概略=あらすじをのべること」と「表面的になでる」ことは違うことに気づいて欲しい。「最も大切なことに的を絞って簡潔に述べる」のが概略だ。卒業生H君の例をあげよう。

パケットフィルタリングとは、最も一般的かつ簡便なセキュリティ技術として知られており、最近のルータは大半が持っている機能である。パケットフィルタリングは、ネットワーク機器に一定の負荷を与えるが、これはパケット転送の遅延を引き起こし、サービス品質の低下につながる。本稿では、田中ら [1] のパケットフィルタリングのモデルを用い、ネットワーク機器におけるフィルタリング負荷を軽減する手順を提案する。

これを読む人は、以下のような疑問を感じるだろう。

- パケットフィルタリングって「どんな」こと?

- 一定の負荷って「どんな」負荷？
- なぜそういう問題が起きるの？
- サービス品質が低下すると「どうなる」の？
- 「何」に適用する「どんな」手順なの？
- 田中 [1] を読まなきゃならわからないなんてひどい。
- 実際、「どれくらい」有効なものなの？

これらのことを、全部きちんと盛り込むようにしなくてはならない。一方で、パケットフィルタリングが「最も一般的かつ簡便」なことや「最近のルータの大半が持っている機能である」ことは研究の内容には無関係だ。「簡潔に書く」とは、このような不要な表現を省くということだ。例えば以下のように書きたい。

インターネット上に存在するルータや L3 スイッチなどのネットワーク機器には、不正アクセスを排除するためにパケットの転送を制限するパケットフィルタリングとよばれる機能がある。フィルタリングはパケットに関するルールの集合を到着パケットに適用することによって実現されるが、ルール数が増大するとネットワーク機器内で転送の可否の判断に要する時間が増加し、通信の遅延を引き起こす。通信の遅延は、音声、映像など実時間サービスの品質低下につながる。本稿では、既存のルールのうち、デフォルトルールを分割して上位に配置することで、フィルタリング負荷を軽減する方法を提案する。分割したルールを上位に移動する際、既存のルールで与えられる評価型が変化する場合がある。これを防止するために、移動したルールの上位に例外パケットのためのルールを配置する。ここでは、例外がある場合、ない場合それぞれについて、分割により負荷が軽減するための必要十分条件を与え、これを用いたルールの再構成アルゴリズムを示す。ルール分割の一例について、負荷が 86% に軽減できることを示す。

先に示した疑問への答えが全て盛り込まれているのが分かるだろうか。これくらい書けば、読む人は「自分にとって、本論まで読み進んでおくべき研究かどうか」が判断できる。

はじめに

「概略」と「まとめ」を読んで論文を読む必要性を感じた人が次に読むのが「はじめに」だ。「はじめに」の役割は「概略」で述べた内容をさらに詳細に語ることで、論文の全体構成を読者に知らせることだ。学術論文には目次がつかないので、論文の構成を読者に知らせる場所はここしかない。

1. 研究の背景と問題提起

できるだけ大きなストーリーを描くことから始めて、研究の必要性に結びつける。着目した分野でどんなことが問題になっているか、それに対するこれまでの研究者の取り組みの経緯を、他の研究論文を引用しながら具体的に説明する。そして、全体の流れの中で、自分の研究がどの場所に位置付けられるのかを明示する。もし、全く新しい研究テーマに着目したなら、問題意識の端緒から書きはじめればよい。

2. 研究の狙い

背景で示した経緯に対応する形で、ここで展開しようとする理論の狙いを述べる。

3. 論文構成の紹介

論文の章立てを示しながら、後の章で数式や記号を用いて説明している内容を、それらを用いずに言葉で読者に伝える。その際、なんのためにそれをするのか、どのような点に着目してそれを行なったのかを記述する。実験がある場合は、実験の結果を簡潔に述べる。

卒業性 G 君の例をあげよう。

帰納推論などの言語学習の研究では、あらかじめ解となる文法が用意され、ネットワークの学習の結果どれだけその文法を忠実に再現できているか、すなわちその文法に合致する入力をどれだけ正確に受理し合致しない入力をどれだけ正確に破棄できるか、ということが主眼とされてきた。しかし、人の言語学習では、そこまで明確な解（文法）があるわけではなく、学習環境の違いにより、人によって違った形の文法が学習されているのではないかと考えた。したがって、ここではその考え方をもとに学習法を提案したい。

日記ではないから、「～と考えた。」はまずい。それはともかく、できるだけ大きなストーリーを描こうとすれば、いきなり帰納推論というのは唐突だろう。言語学習は、いろいろな学問分野で取り組まれているのだから、そのあたりを切口にして、徐々に内容を絞りこんでいく。

人の言語学習がどのような仕組みに基づいて行なわれているのか、という問いは、心理学、認知科学、人工知能など、さまざまな学問領域で取り扱われてきた問題である。チョムスキーが提案した生成文法理論 [1] は、情報科学の枠の中で言語学習を論ずることを可能にした。Gold[2] は、ある言語に属している記号列と属していない記号列が与えられたとき、その言語を規定する文法を求めるといふ帰納推論の問題として、言語学習の問題を形式化した。ある言語に属している記号列は正例、属さない記号列は負例とよばれる。正例と負例の集合に合致する文法を求めるといふ問題は NP-困難であることが知られている [3]。正例からの学習については、多項式時間で文法を同定できるための文法に関する十分条件が知られている [3]。

帰納推論の言語学習が想定する枠組では、文法を規定する語は所与である。また、正例が与えられる順序は任意である。しかし、人の言語学習においては、文法の学習と語の学習は同時に行なわれている。環境から与えられる文は例の一種と考えることができるが、学習者は、正例と負例を区別することはできず、むしろ既存の知識で意味内容を把握できる文のみを正例として取り込むことが行なわれている。これらのことは、人の言語学習における効率や精度に深く関わっている。

大きなストーリー、これまでの研究の経緯、文献の引用、問題提起が盛り込まれているのがわかると思う。もちろん、こんな文章がサラサラ書けるのは、専門の研究者だけである。諸君は、いろいろな文献を読んだり、教科書を参考にしながら、これを書いていくわけだ⁴。

研究の狙いは、「はじめに」の山場になる。比較的書きやすいはずだ。上述した問題提起に対応づけて、自分がここで展開していく内容を述べる。

本稿では、このような人の言語学習の特性に着目し、語の獲得を文法の獲得と同時に進める新たな言語学習の枠組を提案する。学習者は、既存の知識から意味内容を把握できる文のみを正例として学習対象にするとの前提にたち、既存の文法知識との差分がたかだか一語となるような文を例として順に採り入れていく。この意味で、学習者が採り入れる例は、単純な文から複雑な文へと向かう順になる。

ここまで読んだ読者は「何が問題になっていて」「それをどのように解決しようとしているか」がおおまかに理解できる。あとは、順に、各章の内容の概略をのべればよい。

卒業生 H 君の論文の章立てから構成を書いてみよう。

はじめに、田中ら [1] の提案したパケットフィルタリングのモデルを概説する。転送の可否に関する各種のルールを 1 つの論理式として単純化し、アクセスリストを論理式の集合として表す。ネットワーク機器の負荷をルール集合の関数として定義する。フィルタの仕様を、各パケットの条件とフィルタが付すべき評価型の組として定義する。2 章では、負荷を最小化しながらアクセスリストを構成する 3 つの方法を提案し、それぞれの方法によって構成されるアクセスリストの負荷が等しくなることを示す。また、条件式が 3 ビット以下の場合について、これらの方法が負荷最小のアクセスリストを与えることを示す。3 章では、研究のまとめを行ない、残された課題について議論する。

記号や数式を用いず、内容を言葉で表現する様子がわかるだろうか。

準備

ここでは、後の内容を理解する上で必要十分な予備知識を読者に与える。多くの場合、それは用語と記号の定義になる。淡々と正確に定義を述べよう。ただし、読者が知りたいのは単なる定義ではなく、なぜそのような用語・記号が必要なのか、なぜそのように定義してよいのかという、定義の背景や意味、とりわけ根拠である。決めればよいというものではないわけだ。だから、以下のような定義の仕方を常に心がけよう。

⁴実は、この文章の最後の部分は不完全である。「深く関わっている」と主張する根拠が示されていないからだ。どのようなことが根拠になるか、考えてみよう。

ルールは R_1 から順に適用されるので、 $|R_i(F)|$ は R_i 以前のすべてのルールに依存する。そこで、1つのルールを適用することによるネットワーク機器の負荷を1とし、一組みのルール R によって決まるフィルタリングの負荷 $L(R, F)$ を、 $|R_i(F)|$ の重み付の総和として以下のように定義する。

定義 1 (フィルタリング負荷)

$$L(R, F) = \sum_{i=1}^n i |R_i(F)| \quad (1)$$

前後の文脈がないとわかりづらいかもしれないが、 $L(R, F)$ を (1) のように定義する根拠があらかじめ示され、接続詞「そこで」によって帰結である定義へ結び付いている。

また、内容を理解する上でポイントになる概念は、必ず例を交えて説明する。例えば、以下のようだ。

アドレスは、マスク記号 $-$ を伴った論理式とする。送信元あるいは送り先、ポート番号、プロトコルといった条件を、ビット列としてアドレスの上位に付加し1つの論理式として扱う。例えば、ポート番号のビット長を2、アドレスのビット長を8とすると、

$$R_3^A = 0111 - - - - - - \quad (2)$$

は、フィルタリングルールの3番目のルールで、アドレスの上位2ビットが11、ポート番号が1であるようなパケットの転送を許可することをあらわす。

予備知識を与える上で注意しなくてはならないのは読者層だ。読者層をどのレベルに設定するかは、論文の性質によって変化する。卒論の場合、想定する読者は学部新4年生だから、3年次までに講義で習得した基本的な概念の説明は不要だ。

本論

「はじめに」と「準備」では、章を更に節に分割する必要は通常あまりないが、本論では、内容毎にいくつか区切る必要が生じる。章のはじめで、各節の内容に触れながら構成を示す。書き方は「はじめに」の中の、章立ての説明に準ずる。

本論で記述する内容は、各自が得た研究結果とそれを導いた道筋だ。関係式を導いたなら導出の過程が道筋にあたるが、手法を提案するなら、問題のどの部分に着眼したのかがポイントになり、実験を行なうなら実験の条件をなぜそのように設定したのかがポイントの一部になるという具合に、内容によって道筋は千差万別だ。読者が著者のたどった思考過程を1つ1つトレースできることが大切なので、論理の飛躍がないよう丹念に記述したい。

個々の注意点は、プレゼンテーションについて述べた点と変わらない。

- 話の流れが分かるように説明する。
- できるだけ図を挿入する。
- 式の導出を示すなら、導出の手順のポイントを言葉で記述する。

議論

まとめの最後に新たな節を設け、何が可能になったのか、何がわかったのか、問題となった点、やり残した課題などをはっきり書く。

まとめ

研究全体のまとめを行なう。提起した問題、与えた結果について簡単に触れ、ここでやり残した問題や今後検討すべき課題をできるだけ多く深くかかえて、それに取り組む際のアイデアや注意点を述べる。不完全な内容でかまわないので、自由に創造力を発揮して書いて欲しい。これらは、各自の思考の流れを、諸君の研究を発展的に受け継ぐ新4年生に受け渡す大切な情報でもある。

論文誌に掲載される論文の場合は紙面が制限されるので、「議論」と「まとめ」はひとまとめにされることが多い。

参考文献リスト

学会論文誌の場合、参考文献リストには「論文の中で実際に引用した文献」だけを列挙することに

なっている。引用していない文献をリストに含めることはしてはいけない。しかし、卒業論文は研究室の次のメンバーへ研究の成果を伝達するという意味合いが強いので、参考文献リストには、各自が研究を行なう中で参考にした論文、教科書などをすべてのせるべきだ。

文献リストには、学会誌ごとに決まった書式がある⁵が、内容はほぼ同一だ。電子情報通信学会の場合、著者名、論文名、論文誌名、巻、号、引用したページ、発行年の順になる。書籍の場合は、著者名、(訳者名)、書籍名、出版社名、引用したページ、発行年となる。詳しくは、この文書の参考文献リストをみて欲しい。

参考文献

- [1] 田中 賢, 伊藤 聖, "ネットワーク機器の負荷を軽減するフィルタリングルール再構成法", 電子情報通信学会論文誌 (B), Vol.J88-B, No.5, pp.905-912, 2005.
- [2] 松金輝久, 武永康彦, "組合せ最適化問題としてのぷよぷよの連鎖数判定問題", 電子情報通信学会論文誌, VOL.J89-D, No.3, pp.405-413, 2006.
- [3] 木下是雄, "理科系の作文技術", 中公新書.
- [4] 白楽ロックビル, "博士号 とる？とらない？徹底大検証！", 羊土社, 2000.

A 英語論文と輪講

ある分野について効率的に知識を蓄えるには、英語の論文を読むのがよい。英語の論文を読むと、

- 国外の最新の情報を邦訳を待たずに入手できる。
- 基本的な専門用語を原語とともに覚えることができる。
- 英語論文を書く際に必要な表現を会得できる。
- 内容を理解するために必要な教科書・辞典類を参照することで、具体例とともに必要な概念・手法を理解できる。
- 論文の書き方 = スタイルを会得できる。

という具合に、研究に必要な様々な知識を一度に得ることができる。何をやったらいいかわからないときには、英語の論文を読んでみるのは1つの方法だ。

初学者が英語文献を読む場合、最初の1通目を読破することは相当に骨の折れる作業になる。しかし、2通目は1通目の半分くらいの労力で読むことができる。専門用語を含む文の表現は文学作品などに比べればずっと限られたものだからである。この傾向は読めば読むほど顕著になる。関連する文献を4通ほど読めば、ひとおりの知識を得ることができたと考えてよい。

気をつけなければならないのは、いいかげんな文献はいくら読んでも害になるばかりで得にはならないということだ。良い文献とは、対象とする問題や目的、手法が誰にでも納得できる形で簡潔に整理されている文献となるのか。一言で言えば、曖昧な所がないということである。一旦、いいかげんな表現や曖昧な議論に毒されてしまうと、それを正すのは容易ではない。明らかに著者が手抜きをしていると思われるような文献は、早い段階で放棄したほうが身のためである。

A.1 英語論文の読み方

英語文献の読み方をかいつまんで説明するが、英語の読解に関することを除けばそのまま日本語文献にも当てはまることばかりだ。

論文を読むときは通常、

1. 表題 (title)
2. 概略 (abstract)
3. 結論 (conclusion)
4. はじめに (introduction)
5. それ以外の部分

⁵BiBTeX を使って、それぞれの学会の書式に従ったリストを生成できる。

という順番に読む。最初から順に読んで行くのではなく、トップダウンに内容を把握していくわけである。これは、自分の研究に役に立つ内容がどうかを、できるだけ短時間で効率的に判断するためである。

専門的な文献の場合、それぞれの段階をきっちりと理解できていないと、次の段階に進むのは難しいが、英語の場合はその傾向が顕著である。1つの誤訳が、後々まで理解を妨げる、ということが間々あるので、それぞれの段階で疑問点を最小限にしておくことが重要である。

さて、通常は次のようなステップで英文を読み進めよう。

1. 1段落分を読み進める。
2. 辞書を索いて意味の不明な単語や表現を調べる。
3. 再度読む。

この段階で、7割程度は理解できるものだ。問題は残った3割である。意味がとれない原因は、以下のようになっているはずだ。

1. 辞書で索いた単語の意味が不適切。

辞書は、万人に役に立つよう最も一般的な訳から順番に掲載している。ところが、論文とは専門的な内容の文書だから、1番目や2番目の訳語をつけてみても意味が通らない場合が多い。どうもおかしいと思ったら、簡単な単語であっても(簡単な単語こそ)辞書を索いてみて後ろの方まで読んでみる必要がある。数学用語、科学用語という欄に適切な訳が見つかることがかなりある。

専門の論文を読むのであるから、受験用の英和中辞典程度で太刀打ちできるわけがない。使いなれているからといって、このような辞書を使っていると、適切な訳語が掲載されていないから、意味がわからないし、結局辞書を索く時間が無駄になってしまう。最低限これくらいの辞書を使うべきだ、ということで、私は、

「リーダーズ英和辞典」研究社

を以前から勧めている。研究室にはこの辞書のCD-ROM版もある。emacsから使うことができるし、ノートパソコンにインストールして、マメに索きながら読んでもよい。とにかく、わからない英文を理解するためには、まともな辞書を使ってちゃんと索くこと、これが第一歩である。

ネットで使える辞書の定番は、<http://www.goo.ne.jp/>だと思うが、よくできているのは、<http://www.alc.co.jp/>だと思う。翻訳家のグループがボランティアで訳語の登録を続けている英和・和英辞書で、収録語数は100万を越える。例えば、未定乗数法を引くと”Lagrange’s method of undetermined multipliers”がヒットするし、”context-sensitive language”を引くと「文脈依存言語」がヒットする。こういう辞書はなかなかない。

2. 構文がとれない。

英語文献では、3行以上にわたって続くような長い英文も多い。構文を把握することが英文理解の前提になるのは、受験英語と一緒にある。そんな話を始めると本が一冊できてしまうので、ここでは、英語を読む際の手がかりをいくつかあげてみる。

英語の文には必ず「主語」と「動詞」がある。他の部分を全てマスク(隠す)しても、「その部分だけで完全な英文が構成できる」もの、それが主語と動詞である。動詞の語尾のsが主語の人称と一致しているとか、助動詞と動詞の時制に一貫性があるなど、文法上矛盾がない、ということに加え、それだけで意味が通る、ということが重要な点になる。主語と動詞を見つけること、これがまず先決である。

主語と動詞が見つかったら、「文法的に無矛盾でかつ確実に意味が通る部分」を最大化するように部分的に訳をつけていく。たとえば、関係代名詞があったら、その関係代名詞がどの語をさすか、候補がいくつかあるはずだ。その中で確実に意味が通るものを選択していくのである。このとき、「文法的にこれなら正しい」「これなら意味が通る」というだけではなく、「これだと意味が通らないからこうなるはずだ」という考え方が役に立つ。この段階の解釈は、文の意味を考え合わせてはじめて成功するものだ。つまり、闇雲に訳していくだけではダメで、意味を吟味することがどうしても必要なのである(完全な構文解析は意味解析をとまなう、という事実は、自然言語処理の困難の1つになるというわけである。)

確実に意味が通る部分を最大化して行くと、最後はどんな英文もいわゆる5文型に落ちてくる。こうなれば、全体に訳をつけることも簡単にできる。

どうしてもわからない部分が残ってしまう場合、大概は文の切目の目星を誤っている。そんなときは、原点にかえる。主語と動詞の決定を勘違いしていないか。名詞を動詞と勘違いして訳していないか、その逆はないか。そして、もう一度辞書を索いてみることだ。

3. それでも意味がわからない。

英訳は間違いのないのだが、どうも意味がよくわからない、というのはよくあることである。そんなときは、気にせず読み進める方がかえってよい。仮に日本語で表現されていたとしても、1つの文だけで内容を完全に理解することは難しいのが普通である。わからない部分に赤線を索いて読み進み、あとで再度読み返してみることだ。

4. 注意すべきこと。

以下のような読み方は、効率がよいようで実はちっとも役に立たない。

- (a) わからない単語に適当な訳をつける。
- (b) 内容はわからないがまず全文を直訳してみる。
- (c) 最後に訳文を読み返して内容を理解する。

このようなやり方をすると、「わかる部分はわかるのだが、わからない部分はさっぱりわからない」ということになってしまう。なぜそうなるかは、上記の説明で理解できると思う。構文を把握する段階で文の部分的な意味から考えていかなければいけないし、場合によっては単語の意味から考え直さなければいけないというのが現実なのである。それを強引に訳して意味不明の日本語を羅列してみても、わからない部分はやっぱりわからず、結局原文を再度読み直すしかない。それどころか、一度全訳をつけてしまうとその訳にとらわれてしまって(頼ろうとする、といってもいいが)、原文をおろそかにしてしまうことにすらなるのである。

専門的な内容の文献が相手なので、前の部分の理解が後々の理解にまで影響してくるのは、日本語の教科書と変わらない。英文は少しずつ確実に読むのが結局は最も効率的だ。

A.2 英文輪講資料の作り方

英文輪講のやり方は毎年異なるのだが、1人1つずつ文献を担当することが多い。他の研究室では、1つの文献を何人かで分担することが多い。分厚い洋書を分担して読むような場合には、そのようなやり方をとるしかないわけだ。この場合どうしても担当者から見ると話題が断片的、部分的になってしまって、往々にして本人のためにならないという問題がある。そもそも私は、輪講を通じて、単に洋書を読めるようになるだけでなく、れっきとした論文を読めるようになって欲しいのである。論文が読めれば洋書を読むことはそれほど難しくない。

一人1つずつ文献を読んでいく以上、とにかく与えられた文献に関しては担当者が完璧に理解してくるのが基本である。ちゃんと理解した上で、それを他人に理解もできるような形で説明するのが輪講の目的だ。もとより説明している本人もわかっていないような発表はしてはいけない。輪講資料を作るときには、まずこの点に注意すべきだ。

すでに英文は読み説いているとすると、たとえば輪講資料は以下のようなものを用意すれば良い。

- 全訳
言うまでもなく直訳では話にならない。意識をしたものを用意する。英文を和訳した文章には独特の癖が生ずる。日本人が日本語で書いた文章と感ぜられる位にしなくてはならない。訳文を読み上げてみて、自分は普段こんな表現は使わないな、と思ったら、その訳はまだ不十分なのである。
- 概略
自分に割り当てられた時間は、せいぜいのところ輪講2回位である。時間になると4、5時間。文献のボリュームにもよるが、この時間内にすべての発表を終わらせることはまともにやるとかなり難しい。そこで、あらかじめ概略を用意して全体の流れを説明し、重要な部分から原文にあたって読み解いていく、という手順が必要になる。用意する概略の中では、論文の目的、手法、結果、評価などを、図を引用しながら簡潔に説明する。文献を理解するために必要な概念を、他の教科書や文献から引用することがかなりあるはずなので、そのような情報を概略の中にまとめておく。理解を容易にするために、追加的な図を作るのもよい。重要な語句については、調べた結果をリストにして添付する。
- 原文のコピー
全訳をつける場合、どうしても訳せない部分、訳せても意味がとれない部分が残る場合がある。そのような箇所は、原文と訳文にアンダーラインを索いて、発表の中で議論する際の助けになるようにする。

A.3 輪講の進め方

発表者：

資料がきっちり準備されていれば、発表はさほど難しくない。まず、概略を参考にして論文の骨子を解説する。次に、原文を1パラグラフずつ読み上げる⁶。参加者はこのときはじめて内容に触れることになるので、あまり速く読み上げられると途中でついていけなくなる。読んでいる本人が、後戻りすることなく内容を理解していける程度の速度で読むことが肝要だ。読み上げたら、順に邦訳をつけていく。このときも、準備した邦訳をただ読み上げるのはまずい。参加者の頭の中では、原文中でわからなかった表現や単語の意味を重点的に聞き取っていかうとするから、邦訳に対応する原文の位置が常に把握できる程度の速度で訳をつけていく必要がある。一番よいのは、用意した訳文を見ずに、その場でもう一度訳せばいいのだが、これだと逆に時間がかかりすぎていつまでたっても先に進まないことになる。原文と邦訳を同時に見ながら訳をつけていく、というのが丁度よいようだ。

発表中にどうしてもわからない箇所が出てきたら、参加者に意見を求める。それでもどうしてもなくなったら、うんうん唸って議論が暗礁に乗り上げる前に、発表者が「これは宿題にして来週報告する。」としてしまう方がよい。先を読み進むと、あっさり理解できる、ということがかなりあるからである。それでもだめで、いろいろ考え調査してもわからない場合もある。こんなとき、完全に理解することは難しくても、解釈をつけること位はできるのだから、次回にそれを報告する。仮説をたてる、ということである。そこまで考えを進めておくと、どこかで同じような内容の文献に行き当たったときに、不明だった点があっさり理解できたりするのだ。

参加者：

輪講というと、とかく自分の担当以外は関係ない、という態度の輩が出るのだが、それならば資料を受け取ってさっさと部屋に戻って一人で読んでいればいいだけの話だ。せっかく時間をさいて集まっているのだから、参加者全てが何かしら得るものがなくては意味がない。

なぜ輪講をやるか、つまりなぜ日本語の教科書ではなくて英語の文献を読むかということ、専門分野の最新の文献を、スラスラと読めるようになる必要があるからである。ということは、参加者は発表者の用意した全訳を読むのではなく、ほとんどの場合原文を中心に読んでいなくては意味がないのである。発表者が原文を読んでいるときも邦訳をつけているときも、常に原文に注視し、誤訳や誤解がないか注意する。そして、おかしいと思ったら、発表中でもかまわずコメントをするわけだ。

議論：

輪講は、報告会と違って顔も知らない外国人の研究結果を読み解いていく作業になる。外国の論文というのは、相当ひどいものから第一級のものまで実にさまざまで、日本の論文誌に比べるととんでもないものが掲載されている場合もまれにあたりする。ところが、多くの日本人は外人コンプレックスが抜けないから、なんとなく英文になっていると正しいように感じてしまうものだ。これではまずい。

輪講は、内容が理解できてそれでよし、というのではなくそこから先、内容の吟味がより重要だ。文献は、常に批判的に読まなくてはいけない。あえて悪く言えば、目を皿のようにして、矛盾点や欠点を指摘しながら読まなくてはいけない。よい点は文献の著者がいくらでも主張しているものだが、欠点は文献の末尾にさらりと触れられている程度だし、著者が気が付いていない欠点が存在するのがむしろ普通なのである。そのような点に気づく能力を身につけるのが、輪講の最終的な目的と考えるべきだ。だから、発表中はただ単に読み解いていくだけではなく、素朴な疑問を抱いたらそれを書き留めておき、あとで必ず発表者に質問しなくてはいけない。部分的にでも理解できていれば、2つや3つはすぐに疑問点や問題点が浮かぶものだ。輪講のときは全員必ず一度は発言する。そうすることで、他人の考えたことを確実に自分のものとしていくことができる。

発表者は、その質問に対して答える義務があるが、だからといって早急に完全な答えを求めるわけではない。わからないものはわからない、という現実はもちろんある。しかし、それ以上に、その疑問点を種にして議論が展開していったら、新たに勉強すべき事柄がみつかったり、解決法の糸口や新しい問題がみつかったりする、ということが何より大事なのだと思う。

B 専門書について

研究に限らず、情報を得る時には極力オリジナルに当るようになるのが、専門家になる際の近道だと思う。この理由を考えるために、理論や技術が普及する流れを追ってみよう。

コンピュータに関連する理論や技術のほとんどは、海外で考案される。インターネット、プログラミング言語、アルゴリズム、コンパイラ、ニューラルネットなど、その割合は9割以上といってもいいかもしれない。例えば、ある理論が世界的に注目されるとする。有用な理論であれば、やがてそれに関する紹介論文や書籍が海外で出版されはじめる。平行して、国内でもその理論を取り入れよう

⁶専門書の場合、英文の読み方がわからないことが多い。例えば、 $\frac{1}{n}$ や $x^2(i+1)$ をどう読むか。

とする研究者が徐々に現れる。注目の度合が高ければ、それらの論文や書籍の邦訳、解説論文、邦訳を元にした教科書が出版されるようになる⁷。このように、我々に届く情報の多くは翻訳や解釈を何度も経ることで、オリジナルのものからはどんどん変形してきたものだ。

解釈を経た情報は、一見わかりやすくまとめられているが、弊害として「その概念が何を表そうとしたものか」とか、「なぜそのようなことを考えたのか」という部分の情報が欠落してしまっていることが多い。結果を導くための手順ばかりが重視されて、そのようなことを考える背景や仕組みが理解できないことになってしまう。定期試験まで2日を切った状況では背景も仕組みもないけれど、時間をかけて新しい研究テーマに取り組もうとするなら、むしろその「なぜ」「何を」の部分を知ることが後々重要になって来るように思う。和文よりも英文、教科書よりも論文、論文なら原著論文にあたるよう心がけたい。

教科書を選ぶ場合にはどうしたらよいただろう。英語の書籍を読むのは骨がおれるなら、できるだけオリジナルに近い質の良い邦訳を選ぶのが吉だと思う。一般論だが、わざわざ邦訳されるほどの本は内容も定評のあるものが多い。そのような本も含めて、かねてからコンピュータ関連の良書のリストを作っておきたいと思っていたら、昨年、

「コンピュータの名著・古典100冊」石田晴久 編・著 インプレス

という本を見つけた。異論はあるものの、概ね納得できる内容だったので、100冊のうち今でも入手できるものをまとめて購入したいと思っている。

C 博士前期課程への進学について

情報科学科の場合、学部卒業生のうち10人程度が大学院博士前期課程に進学する。しかし、就職してみれば分かるが、情報系で修士持ちは世間では半ばデフォルトである。国公立で半分、私立でも中堅以上の大学は3割以上は進学していると思う。これには理由がある。

1. 企業へ就職する場合、博士前期課程修了者は学部卒よりも採用面で有利。職種によっては、学部卒を採用しない。
2. 修士は、就職が2年遅れて学費を2年余計に払うことになるが、初任給や昇給体系が学部と違うので、生涯賃金で考えると学部卒より多くなる。

実際のところ、「もっと勉強したいから修士に進みたい」という学生はあまりいなくて(本当はそうあって欲しいけれど)、もっと現実的・実利的な理由で進学が選ばれている。もう2年在籍して将来をよく見定めよう、とか、自分の世界を広めておこう、とか、動機はいろいろだ。学部と比べると単位の取得が容易で自由な時間が多いから、その分いろいろな過ごし方が考えられる。

昨今は就職戦線が厳しさを増しているのは承知の通りで、学部卒という肩書がハンデになってしまいうケースは結構あると思う。神奈川大でも、22年度から奨学金や学費の減免の制度が大幅に拡充され、概ね半分程の学費で就学できるので、事情が許す限り進学することを勧めたい。

D 博士後期課程への進学について

博士前期課程に比べ、博士後期課程は学生から見るととかく謎が多い。研究を面白いと感じる学生ならば、有意義な生活が送れるに違いない時間なのだが、何のために進学し、何をもちて修了となるのか、修了後はどうなるのか、という基本的な情報が表立って語られないため、敬遠されている感がある。ここでは、博士後期課程の実態について、少しだけ説明しよう。

D.1 何のために進学するか

大学を卒業すると学士(Bachelor, バチェラーと読む)という学位が大学から授与される。諸君の場合、付与される学位は理学士(Bachelor of Science)だ。大学卒業の後、博士前期課程を修了すれば理学修士(Master of Science)を授与される。そして、さらに博士後期課程を修了すると授与される学位が理学博士(Doctor of Science)だ。博士課程に進学するのは博士の学位を取るためだ。大学で

⁷邦訳は、学生諸君が思っている以上にいいかげんだ。情報系学生の座右の書、K & Rの邦訳は誤訳が多いことで有名である。邦訳は研究者が行うことが多いが、実際には大学院の学生と共同で行った英文輪講の資料がベースになっていたりする。研究者単独で訳す場合でも、解釈そのものが誤っていることも多いし、新しい概念に当てはめた訳語が不適切で理解を妨げることも多い。

学位といった場合、通常は博士の学位のことを指すので、以下では博士の学位を略して学位ぶことにしよう。

さて、何のために学位を取るのか。それは、研究者として飯を食うためである。もちろん、学位を持たずに研究者として活動している人はいるし、企業で開発に携わる技術者で学位をもっている人もたくさんいる。しかし、大卒の資格すなわち学士をもっていなければ一流業への就職が難しいのと同じように、学位をもっていなければ大学や研究機関ではほぼ完全に採用してもらえないという現実がある。学位は、研究者の免許証のようなものなのだ。

では、研究者を目指さない人にとって、学位をとることの利点は何だろうか。企業が学位保持者を採用する場合、その人に期待するのは、もはや駒働きとしての役割ではなく、自力で1つの問題を解決していく責任者としての役割だ。だから、学位保持者は最初からある程度まとまった仕事を任せられ、給与についても修士卒よりは高くなる企業が多い。社内の中でも周囲から一目置かれることになる。企業に入って、人に言われるままに仕事をこなすことに飽きたらなければ、学位を取って専門分野を究めてから就職するのは悪くない。

研究者を目指さないといっても、企業に入って開発作業に携わるなかで、自分の進む道を再度見極めたいと感じることもあるだろう。そんなとき、学位を持っていれば手持ちの選択肢が増えるはずだ。今は学位が必要なくても、将来役立つことはありうるわけだ。

D.2 修了の条件

博士課程修了の条件は専攻ごとに異なるが、情報科学専攻の場合、査読付原著論文を2本書くことが修了の必要十分条件になる。2本のうち最初の1本は博士前期課程での研究結果でも差し支えない。また最近はこの1本に減らす方向のようだ。博士課程の在籍年数は3年、上限は6年だ。3章で紹介した論文誌に投稿することが一般的だが、それが必須というわけではない。査読の緩い論文誌にだして、早めに数を確保しようとする人もいる。意外と簡単だと思うかも知れないが、論文を投稿してから査読を経て採録になるまでには半年程度の時間を要するし、散々時間をかけて不採録になることもある。この辺りが、博士後期課程の修了を不透明にしている一因だ。

D.3 修了後のこと

修了後、もし企業に就職する道を選ぶなら、修士卒と同じように就職活動をすることになる。研究職へ就くことを希望するなら、大学や研究機関の公募に応募することになる。公募の情報は、それぞれの分野の学会が毎月発行する学会誌から得ることができる他、独立行政法人科学技術振興機構が運営するJREC-IN(Japan REsearch Carrer Information Network)のwebページが広い範囲を網羅しているので役立つ。実際には、純粋な公募とは異なるプロセスで決まっていくポストも多い。例えば、博士課程の学生が、そのままその研究室のスタッフとして就職するのはたまに見かける光景だ。この辺りは、一概にこうといえず、この業界の微妙な部分なのだが、進学を考えている学生諸君にとっては重要なことだと思うので、進学を決める前に私に直接話を聞いて欲しい。

博士課程に進むデメリットも触れておかなければ不公平だろう。企業に就職しようと思っても、博士課程修了者は受け入れないという企業も多い。学位保持者は、企業から見ると完成品であって、他の修士卒の学生と比べ色をつけずらい、つまり思うように育てずらいという事情があるようだ。研究所を持つような大企業が、完成品としての学位保持者を採用することは多いが、長引く不景気で企業の研究所はどこも基礎研究部門を縮小している。企業への就職を考えるなら、意識的に企業が欲しいと思うような研究テーマを手掛けるのは、1つの手かも知れない。

博士号の取得を考える際、[4]はためになる。ただし、この類の話は研究領域や大学によって異なる部分が多いので、あくまで参考程度にとどめたい。

D.4 奨学金について

前期課程以上に、後期には様々な奨学金が用意されている。大学は博士課程の学生を奨励しているので、大学が用意する奨学金を得られる可能性はかなり高い。それ以外にも、日本学生支援機構の奨学金は一般に博士課程の学生には優先的に配分される。また、修了後一定期間研究教育機関に勤めれば返済を免除される(されました)。

D.5 他大学への進学について

毎年何人か他大学の院を受験する学生がいる。僕自身学部と院は別大学なのだが、それだけにこの選択のいい面と悪い面はよく分かっている。もしそういう考えをもっている学生がいたら、個別に相談にのるので部屋を訪ねてほしい。

E 研究テーマの作り方

小学校の頃、「夏休みの自由研究」というのが毎年あった。なにかを「研究」してその成果を休み明けに報告せよ、ということなのだが、はて、何をしたら研究になるのか、当時の自分にはよくわからなかった。今から思うと、当の先生自身、何をやったら研究になるのか、分かっていなかったのだと思う。

研究とは、対象に疑問を抱き、その疑問を解消することだ。対象はなんでもよくて、人間の行動だったり社会の仕組みだったり自然現象だったり。大切なことはその疑問を「まだ誰も抱いたことがなくて」「自分がはじめてその問題に取り組む」ということだ。そんなことが全ての小学生にできるとはとても思えないが、それはさておき。

研究とはそういうものだから、あらかじめ答えが用意されていないのは当然のことだ。未知の対象を探るといっても、そんな訓練は受けたこともないし、そもそもどんなことを言っているのかよくわからないかもしれない。以下では、新しい研究テーマを見つけるプロセスから、研究の端緒を得る方法を示す。

E.1 深く知り欲を持ち疑問を抱く

研究をするには、まず研究テーマをはっきりさせなければならない。研究テーマとは例えば以下のようなものだ。

例) フィルタリングルール最適化問題 (田中研)[1]

パケットを転送するか否かをルールの集合によって判断するパケットフィルタリングというルータの機能がある。ルールの書き方が不適切だと、フィルタリングの効率が落ちて遅延が発生してしまう。遅延を減らすようなルールはどのように構成すればよいだろうか？

この研究を始めたきっかけは以下のようなようだった。あるとき、工学部の Firewall のフィルタを書く必要があった。何度か書きなおしていると、「はて、本当にこれが一番よい書き方なのだろうか？」と疑問に思うようになった。一番よい書き方を知りたいと思ったのである。じゃ、「一番よいって何？」と考える。「そうか、“良さ”を定義すればこれは研究テーマになるな。」と気付いた。

拍子抜けするほど単純なことなのだが、研究テーマを得るきっかけはこんなものだ。ただし、単純ではあるが1つのことに深く取り組むことが大切なのだと思う。浅くではいけない。深くである。

深く取り組める対象という、諸君の身近な所ではゲームやインターネットがあるだろう。1つのゲームに徹底的に取り組むとき、何らかの疑問が湧かないだろうか。「これはもっとこうすればいいんじゃないの？」とか「これはおかしいんじゃないの？」という気持である。対象は違うが、研究テーマの端緒は、そのような「疑問」や「提案」から始まるのだと思う⁸。

研究テーマを得ようとするとき、大切なことはこの疑問や提案を意識して抱くようにすることだ。ただ漫然と取り組むのではなく、意識して疑問を絶やさないようにする。悪くいえば、イチャモンを付けるのだ⁹。

直観的な疑問を上手に捉え、形を与えてやれば、それがそのまま問題になる。学生諸君が学部の講義で教わったコンピュータに関連するいろいろな学問分野、プログラミング、コンパイル、人口知能、などなど、いずれも先人の「こうしたい」という思いが最初にあって出来上がってきたものだし、その追及は今現在も続いている¹⁰。研究とは、その「こうしたい」の一番先端の部分に参加すること

⁸ 手元の電子情報通信学会論文誌を見ると、「組合せ最適化問題としてのぶよぶよの連鎖数判定問題」[2] という題名が目にとまった。ぶよぶよだって立派なテーマになるのだ。

⁹ 疑問のベースにあるのは、つまるところ「こうあってほしい」という欲かもしれない。面倒くさがりや人はプログラマに向く、という話がある。面倒くさがりやの人は、なんとか手を抜こうとしてこまめにアルゴリズムを考えプログラムを書く、というわけだ。欲がまったくない人は善人かもしれないが進歩しないね。

¹⁰ 普通、人は自分が興味があるものにしか徹底的に取り組まないと思いがちだが、実はこれは鶏と卵の関係で、深く取り組みれば取り組むほど対象について深く知るようになって興味が湧いて来るものだ。自分がかつとも関心がなかった講義でさえ、その内容を知れば自然と関心がわくものだ。多少の辛抱さえ厭わなければ、世の中の多くのことが興味深い対象であることに気付く。

だといっていい。欲と疑問をもって先端の部分に深く関わっていけば、研究テーマは自ずと見えて来るものだと思う。

演習問題

自分が最も深く関心を抱いた事柄を思い浮かべ、疑問点を5つ以上書き出してみよ。

F 科学的思考と技術的思考

科学や技術に関する問題にはそれぞれ共通の取り組み方がある。というより、そのような取り組み方をするのが科学的、技術的なことだ、と言ったほうがよいのかもしれない。

サイエンスすなわち科学が目指すことは、「真理の探求」である。有用性があるか否かにかかわらず、真理が明らかになることに価値をおくというのが科学の立場だ。例えば、ある問題が解けないことが証明されることに通常の意味での有用性はないけれど、科学としては価値がある。

科学的方法とは以下のような方法である。

科学的思考

1. 対象を単純化して表現する
2. 問を表現する
3. 問の解を求める

数学は対象を表現する方法を考える学問だ¹¹。例えば、画像を表すにはベクトルを用いればよいし、論理を扱うなら代数系を用いればよい。学生諸君が大学で数学を教わるのは、対象を表現する方法を知るためだ。

物理学は自然を表現する方法を探す学問だ。1つの表現法を設定して、対象の動作を予測する。つまり、対象の動作がどうなるか、という問をその表現法にもとづいて表し、解を求める。解が実際の現象とどこまで一致するか、その一致度をどこまでも高めていこうとするのが物理学だ。

コンピュータの分野にも「科学」がある。諸君が所属するのは情報科学科で、実際情報科学という学問分野は存在する。しかし、実のところ情報科学が何を指すのかは漠然として今ひとつはっきりしない。「情報」といってしまうとおよそどんな学問分野でも情報を扱うから、なんでも情報科学ということになってしまうわけだ。

コンピュータで情報を扱うことを前提にすると、コンピュータが何ができるか、どうすればよいか=どのようなアルゴリズムを使えばよいか、というように、問題がはっきりする。このような立場から情報を捉える学問分野をコンピュータサイエンスという¹²。コンピュータにまつわる問題を数学的にきちんと取り扱って、計算における種々の真理を明らかにしていこう、という狙いが込められているわけだ。

諸君が所属するのは理学部情報科学科だから、本当は科学を目指す学科なわけで、プログラムをどのように作るか、とか、サーバをどう運用するかといったことは最終的には二の次である。現実には、具体的なコンピュータに関与しなければ何も問題は埋まれてこないし、就職を考えたらそう科学科学とばかりも言っていられないわけだけど、本当はこういうことなんだ、というのは知っておいてほしい。

では、技術的思考は何を目指すか。それは単純で、「もっと便利にしよう」ということだ。もっと速く、もっと大量に、もっと簡単に、それがいつの時代においても技術=工学(engineering)の目指す所だ。便利にしようとするときに科学で得られた知見や手法を応用するとうまくいくことが多い。今日では、技術が複雑になってその傾向がますます顕著になりつつある。だから、科学と技術は今日では地続きで、科学技術と一括りにされることが多い。

科学と技術が地続きであることが分かった所で、研究テーマにどんな型があるかを次に述べる。

¹¹表現する対象がなくてもよい、というのが純粋数学の考え方だ。

¹²海外の大学では、コンピュータサイエンス学科という名称が一般的で、情報科学科とか情報工学科とかいう名称に相当する学科は存在しない。

G 研究テーマの型

疑問を抱けばそれが研究テーマになるといっても、多くの疑問や願望は漠然としたものでそのままではどう取り組んだらいいのかわからないだろう。対象によっては、「どのような疑問を抱けばいいのかわからないかもしれない。ここでは、研究テーマの型を分類してみる。それぞれの型は、そのまま研究テーマの新規性、すなわち既存の研究との差分に対応する。全ての学問分野の型を分類することは難しいけれど、ソフトウェアの主要な研究テーマは網羅できていると思う。

G.1 科学的思考に重点を置いたタイプ

G.1.1 問題定義型と問題解決型

研究テーマを分類する軸はいくつかあるが、もっと大切な軸は問題が既知のものか否かによる分類である。

問題定義・提案型

漠然としか語られていなかった、あるいは着眼されていなかった問題に、数学的な定義を与えるというタイプ。対象は本質を失わない範囲で単純化され、それによって枝葉にとらわれない議論を可能にする。

例) フィルタリングルール最適化問題 (田中研)[1]

(考察の対象)

パケットを転送するか否かをルールの集合によって判断するパケットフィルタリングというルータの機能。

(疑問) ここがポイント

ルールの書き方が不適切だと、フィルタリングの効率が落ちて遅延が発生してしまう。遅延を減らすようなルールはどのように構成すればよいだろう？

(研究結果)

ルールはそのままでは複雑すぎるので、論理式の集合として単純化してしまおう。

問題解決型

やりたいことや目的がはっきりしていて、すでに形式化されており、その解き方を考えるというタイプ。

例) 矩形パッキング問題

(考察の対象)

矩形モジュールを配置する際の面積を最小にするという既知の問題。

(疑問)

既存のやり方よりもっと面積を小さくするには？

(研究結果)

問題定義・提案型は、実際にはその解法と合わせて研究テーマになっている場合が多い。一般には、問題解決型の研究よりも問題定義・提案型の研究の方が大切であると言われる。問題が形式化されていないければ問題を解くことはできないという点で、問題定義・提案型はオリジナリティが高いからである¹³。

問題定義・提案型は、科学的思考の1.と2.に関わるもので、問題解決型は3.に関わるものだということが分かるだろう。ここで述べたようなアルゴリズムに関するテーマだけではなく、漠然とした問題にこのような定義を与えることで、コンピュータで扱えなかったことが扱えるようになることに注意したい。認識、学習、推論、言語など、このような形式化によって新たにコンピュータサイエンスの対象となった問題は数知れない。学生諸君には、是非問題定義・解決型の研究に取り組んで、コンピュータサイエンスの対象をもっと広げてほしい。

¹³ただし、このことは問題解決型の研究が問題定義・提案型の研究よりも容易であるという意味ではない。適切に形式化された問題については、すでに多くの研究結果があり、新しい解法を生み出すのは容易ではないことが多い。

G.1.2 モデル・手法指向

問題を形式化するにはそのための道具が必要である。工学においては数学的な表現手法を用いることになるが、生の数学をもう少し使いやすくした理論を用いると有用であることが多い。問題と数学の中間に位置し、問題やその解法の表現・解析手段を提案するのがモデルである¹⁴。それまでどのように扱ったらよいかよくわからなかった問題が、そのモデルによってすっきりと説明・解析できるようになることを目指すのだ¹⁵。

モデルは数学と同様抽象物なので、それ自体を実世界と切り離して際限なく議論して行くことも可能だ。しかし、工学者は有用性を目指すので、純粋なモデル論を展開することを嫌う。モデルを発展させたり解析したりするなら、それと対応する現象をあらかじめ示せ、というのだ。したがって、工学の研究を行う限り、モデルに軸足を置く場合でも、問題への比重はゼロにしない方がよい。なんのためにそれを行うのかだけは忘れないようにせよ、ということだ。

モデル提案型

問題に取り組むにあたっての有用な概念、表現手法を提供するタイプ。

例) ニューラルネットワーク

人の脳の神経回路網から作られたモデル。パターン認識、制御などの際の学習を表現するのに適する。

例) 遺伝的アルゴリズム (GA)

突然変異と自然淘汰という自然界の適者生存の仕組みをヒントに作られた最適化アルゴリズム。

モデル提案型は、科学的思考の 1. のための表現手段を与えるものだ。モデルに軸足を置くなら、新しい概念を考案しない研究もある。

実験型・解析型

既存の問題や手法について、知見を得るというタイプ。数学的に真偽や有効性を証明することができれば理想だが、それができなければ、十分な量の計算機実験によって普遍性のある結論を得ることを目指す。

例) 遺伝的プログラミングにおける各種初期個体生成手法の比較

(考察の対象)
遺伝的プログラミングの初期個体生成法
(疑問) どれが一番有効かな？
(研究結果)
比較実験を行ってみる。

G.2 技術的思考に重点を置いたタイプ

G.2.1 きちんと定義された問題が全てではない

コンピュータは、手順さえ細かに書いてやれば入力と出力の任意の変換が可能であるように見える¹⁶。だとすれば、必ずしも「問題」として形式化されたものだけを対象とする必要はなからう、便利ならいいじゃないか、との楽観的な立場も成り立つ。

コンピュータが他の機械と最も異なる点はなんだろう。自動車のキーを回せばセルモーターが回るし、マウスをクリックすれば google の web サーバのデータベースに検索がかかる。どんなもので

¹⁴物理学は数学を用いて自然のモデル化を行う学問である。コンピュータサイエンスの場合、例えばオートマトンは「計算」という概念をきちんと考えるために考案されたモデルである。

¹⁵実際には、新しいモデルとそれを用いた解法は同時に提案されることが多い。用いるデータ構造によって計算の効率が劇的に変わることを知っているはずだ。

¹⁶「コンピュータサイエンスとは、すべての問題は中間層を入れることで解決する、という信仰である。」と言った人がいる。

あれ、機械とは作れば作ったように動く。その点で、コンピュータと他の機械とに本質的な差はないように思える。

それでも、コンピュータが他の機械と著しく異なるように見えるのは、ソフトウェアの存在だ。他の機械では、その動作を記述するだけではだめで、動作の仕組みをハードウェアとして実現する方法を考えなければならない。しかし、ソフトウェアを用いれば、外界とのやりとりをデジタル化して行うことで、「外界における要求や問題の解決を、そのためのソフトウェアを作る」という問題に帰着させることができる。これが、コンピュータの応用範囲をほとんど全ての領域にまで広めた原因だ。コンピュータはいわば「万能中間層」なのである。

データ処理型

入力と出力との適切な関係を実現するための中間層を設計するタイプ。

例) 日英翻訳

(考察の対象)
日本語文を英文に翻訳する問題。
(疑問)
自然な英文に翻訳するには？
(研究結果)

コンピュータ上の概念で、この発想から得られたものは数知れない。プロセス間通信、TCP/IP、GUI、オブジェクト指向など。ハードを作るのちがいが、ソフトウェアはその存在を隠蔽することが可能なので、中間層は求められるまま際限なく増加していく。

学生諸君が一番とっつきやすいテーマはソフトウェアの製作だろう。実際の製作が主になるという点で、これはデータ処理型の研究テーマと近い所にある。

製作型

設計し実装するという作業を行うタイプ。身近な問題に関わることが多いので、予備知識がさほどなくても疑問を抱くことが可能だ。効率や精度を追及する製作物なら、既存のソフトウェアとの比較が可能なので研究テーマとしても扱いやすい。

例) FireFox へのタイリング表示機能の実現

(考察の対象)
web ブラウザ
(疑問)
タブだけだと一覧性が低くて使いづらい。タイリングもできたら便利では？
(解法)
作ってみました。

例) 対話モデル Risk Agora における主張の撤回と実装

(考察の対象)
Risk Agora という対話モデル
(疑問) 特になし
(解法)
対話のモデルケースを記述してその動作を実装してみた。

実は、コンピュータはすでに「万能中間層」の段階を越えている。現実の世界とのインターフェースは人とのやりとりのみに限定し、コンピュータそのものの世界をどんどん精密化している。コンピュータの世界のための新しい概念や発想を生み出すのは、白紙のキャンバスに絵を描くようなことで、ある意味では最も創造的な行為かも知れない。

創造型

例) オブジェクト指向プログラミング

(考察の対象)
プログラミング
(疑問) モジュールの独立性と利用可能性をもっと高めるには?
(研究結果)
やりとりはメッセージに限定して外から中が一切見えなければよい。

無から有を描くとはいっても、本当の意味では無からは何も生まれない。恐らく着想はコンピュータ以外の分野から得ることになる。その意味で、このタイプはモデル提案型と似通っているけれど、コンピュータの世界そのものを対象としているのが決定的な違いだ。

G.3 研究テーマ増産のからくり

ここまで、以下の7つの型を示した。

1. 問題定義・提案型
2. 問題解決型
3. モデル提案型
4. 実験型
5. 解析型
6. データ処理型
7. 製作型
8. 創造型

これに加えて、既存の研究テーマから、新しい研究テーマを作るための方法がある。組み合わせ、発展、融合である。これを補助型とよぶことにし、同様に型名を与えよう。

組み合わせ型

ある問題について考えられた解法を、他の問題に適用してみるというタイプ。N個の問題とM個の解法があれば、そこには $N \times M$ 個の研究テーマが潜在的に発生する、という程実際は単純ではないが、新しい研究テーマに詰まったときに最もよくやる手である。

例) 遺伝的アルゴリズムを用いた正規言語同定問題の解法 (田中研)

(考察の対象)
未知の正則言語の例が与えられたとき、それに合致する状態数 n 以下の有限オートマトンを求めるという正則言語同定問題。突然変異と自然淘汰という自然界の適者生存の仕組みをヒントに作られた最適化アルゴリズム。
(疑問)
組み合わせたらもっとうまくいくかな?
(研究結果)
オートマトンをコード化する方法。

発展型

Aという問題を含むA'という問題を新たに自分で提案して解を求めるタイプ。A'という問題を考えることの合理性が研究テーマの価値を決める。問題が発展しているというだけで、これも組み合わせ型の一種ともいえる。

例) RHON による文脈依存言語の認識 (田中研)

(考察の対象)
再帰型高次結合ニューラルネットワークと形式言語
(疑問)
1つ上のクラスの言語だったらどうなるのかな?
(研究結果)
言語の受理系をシミュレートしてみる。

整理・融合型

A というモデルと B というモデルは、実は融合して同じ枠組みで考えられ、これにより双方の利点を生かすことができる、というタイプ。組み合わせ型的一种だが、問題と解法の組合せではなく、モデルや解法を組み合わせる点が異なる。

例) 演繹学習と帰納学習の融合

(考察の対象)

人口知能における機械学習法

(疑問) 人の学習では機能と演繹の両方が交互に起こるのでは？

(研究結果)

融合した学習モデルを提案。

G.4 難易度について

研究テーマの 10 通りの型について述べた。これらを「必要とされる新しい考えの割合」の順に並べると以下ようになる。

1. 問題定義・提案型，モデル提案型，創造型
2. 問題解決型，解析型
3. 実験型，データ処理型
4. 製作型

理論的，抽象的なものから，実践的，具体的なものへと並ぶ。これは，1 つ面からみた研究テーマの難易度の順かもしれない。1. のタイプは，全く新しい問題や概念を作るという点で最も難易度が高い。2. のタイプは，既存の問題や概念を対象とするという点で 1. よりは難易度が低い，競争相手が多く，また問題に対する深い理解と考察が必要という点で，1. 以上に難易度が高い場合も多い。3. のタイプは，やってみれば何かは出るから卒研や修論としては取り組みやすいだろう。工学寄りなので，結果の有効性が評価の鍵を握る。4. も同様だ。補助型の難易度については，基となった研究テーマの型の難易度に準ずると考えてよいだろう¹⁷

研究の型は，学問分野や研究テーマだけでなく，研究者によっても異なる。実験ばかりやる人もいれば解析ばかりという人もいる。中には，自分の型に合わない研究スタイルの存在を認めない人もいて，そのような人は自分の型からの批判ばかり繰り返す。学生諸君が研究室を選択するときには，研究テーマに関心が持てるか否かだけでなく，自分が得意な型を好む指導者かどうかにも気をつけた方がよいと思う。

演習問題

電子情報通信学会論文誌の論文の中から，7 つの型それぞれについて対応する研究テーマを見つけてみよう。

¹⁷ 学問分野の中には，組み合わせ，発展，整理・融合を際限なく繰り返すことが学問の手法そのものになっているような分野がある。そのような分野における創造性とは何だろうか。ノーベル賞を受賞した化学者からしばしば聞かれる言葉は「実験の失敗が成功に繋がった」である。日本人のノーベル賞受賞分野に生理学，医学，化学が多いのには理由があると個人的に思っている。これらの分野は長い地道な努力が求められる分野で，物理学のように天才的な閃きがものをいう世界とは異なる。実に物事を進める日本人の気質が医学や化学の分野に適していること。これに加え，神様があたえてくれた千載一隅のチャンスを見逃さない繊細な観察眼を日本人が備えていることがこれらの分野で大きな成果が上がる主因だと思う。