

(オンライン) ISSN 2432-1036

(冊子版) ISSN 0286-116X

都城工業高等専門学校

研究報告

第56号

令和4(2022)年1月

目 次

研究論文編

- テキストデータの圧縮率を高めるための前処理について……………中村博文・湊田孝康……………1
- Cu₂ SnS₃ 薄膜太陽電池における高効率化の検討……………内村友宏・荒木秀明・中村重之・赤木洋二……………13

資料・紹介論文編

- ワープロ演習での文集作成の一実施例……………中村博文……………19
- 分散登校および遠隔授業による新型コロナウイルス感染症に対する
学内感染リスクの低減効果について……………田中守……………31

研 究 論 文 編

テキストデータの圧縮率を高めるための前処理について

中村 博文¹・淵田 孝康²

Pre-processing for Enhancing Text Data Compression Ratio

NAKAMURA Hirofumi¹ and FUCHIDA Takayasu²

(令和3年10月1日受理)

あらまし バイトデータを対象とした既存のデータ圧縮プログラムの適用前に、前処理として圧縮対象を加工しておくことで、圧縮率を向上できる場合のあることを実験によって示す。本研究において、前処理は、試みとして、「頻繁に現れる隣接バイト対の中で比較的効果的なものについて、すべての出現箇所について2バイトから1バイトへと未使用のバイトコードに置き換える」ということを繰り返している。この方法を活用できる圧縮対象は、バイトデータの中でも、未使用のバイトコードが存在することが確実な、テキスト形式のファイルである。本方法による圧縮率の向上は、既存のデータ圧縮プログラムの圧縮性能が高いと少ないが、ある程度の長さのテキストファイルなら圧縮性能の向上が見られることが多い。gzip、bzip2及びctwを用いた実験において、gzipについては、圧縮率を更に数%縮められる場合があった。本方法は、比較的効果的な隣接バイト対を探すことが必要なため、符号化の時間は桁違いに増すことがあるが、復号で行う前処理の逆の処理は、バイトデータを置き換えながら一度走査する程度の処理が増すだけである。全体として可逆であっても、前処理による置き換えが情報源の性質をかく乱して、情報源モデルによっては表現に必要な平均情報量を増やしてしまうのではないかという疑問について、単純マルコフ情報源という観点では、前処理による置き換えが表現に必要な平均情報量を増やすことはないことを確認する。

キーワード [データ圧縮, 前処理, 隣接記号対, テキストデータ, 文字コード]

1 はじめに

多数のデータ圧縮アルゴリズムが提案されてきている(例えば、文献1~4)。データ圧縮は、圧縮対象に何らかの特徴や構造を見出すために、圧縮対象を解析したり、統計的性質を計量するなど、圧縮対象から得られる何らかの情報をを用いて2値符号化している。なお、念のために明記しておく、圧縮対象が音声や画像に限られるわけではないため、可逆圧縮を前提としている。

データ圧縮アルゴリズムがアプリケーションソフトウェアに組み込まれている場合もあるが、特に

データを選ばないデータ圧縮プログラムやアーカイバは、バイトデータのファイルを対象として提供されている。

本報告では、バイトデータを対象とした既存のデータ圧縮アルゴリズムの入力の前に、圧縮対象を一度加工することで圧縮率を向上させることを試みている。その加工とは、頻繁に現れる隣接バイト対の中で比較的効果的なものを別の未使用のバイトコードに置き換えておくことである。このため、未使用のバイトコードが確実に存在すると考えられるテキストファイルを対象にして実験を行った。

1 都城工業高等専門学校一般科目
2 鹿児島大学大学院理工学研究科

General Education Division, National Institute of Technology(KOSEN), Miyakonojo College
Gradient School of Science and Engineering, Kagoshima University

2 既存符号の符号化と復号の定式化

本報告では、既存のデータ圧縮の符号化を E 、復号を E^{-1} のように表す。また、 E の入力である圧縮対象のデータを I 、 E が I を入力して得られる圧縮結果を C_I と表す。

従来の通常のデータ圧縮法では

入力は I , $I \rightarrow E \rightarrow C_I$, 出力は C_I

とデータ圧縮をして、 C_I を記憶や通信に用い、データを使用するより前に

入力は C_I , $C_I \rightarrow E^{-1} \rightarrow I$, 出力は I

と復号して使用していると捉えることができる。ここで、復号して得られるデータは元のデータ I と同じ内容であることから、復号結果についても記号は圧縮対象のデータと同じ I を用いて表している。

3 提案手法

3.1 提案手法のおおまかな流れ

提案手法はおおまかには、 E に I を直接適用するのではなく、 E に対する前処理 P を付加することを考え、 I に P を施して得た内容 (I^* と表す) を E に適用するというものである。

すなわち、符号化は

入力は I , $I \rightarrow P \rightarrow I^* \rightarrow E \rightarrow C_{I^*}$, 出力は C_{I^*}

と行う。 C_{I^*} を記憶や通信に用い、使用するより前に

入力は C_{I^*} , $C_{I^*} \rightarrow E^{-1} \rightarrow I^* \rightarrow P^{-1} \rightarrow I$, 出力は I と復号して得られた I を使用する。ここで、 P は E を行う前の前処理を、 P^{-1} は P の逆の処理を表す。

この形式は、符号理論において符号器をうまく構成して総合的に通信路をより効率的に使用することと同様の構造である。通信路の効率的な使用は、通信路の性質の理解や定式化が進んでいることから可能になっている。

しかし、データ圧縮において、 E にとって圧縮率の面で最適な I^* がどのようなものであるかや、 I を元に最適な I^* を作り出す処理 (P が行う処理) をどのようにすると処理時間や記憶領域の増加を抑えられるかは、まだ分かっていない。

本研究では、いくらかでも効果のある I^* の作成を試み、前処理による圧縮率改善の可能性を示す。

やや、具体的には、「頻繁に現れる隣接バイト対の中で比較的効果的なものについて、すべての出現箇所について2バイトから1バイトへと未使用のバイトコードに置き換える」ことを必要なだけ繰り返して I^* を得る。

このようにするとき、 I^* は、どの隣接バイト対をどの未使用バイトコードに置き換えたかという情報の蓄積 (I' と表す) と、(I' に記録されている) すべての置き換えを行った後の (置き換えられなかった部分も含む) 内容 (I'' と表す) とから構成される。

つまり、 $I^* = I' + I''$ である。

I' の決定においては、まず I について隣接バイト対の頻度を調べ、頻度の高いものから K 種類まで (足りない場合はその数まで) を置き換える候補と考え、各候補について、圧縮対象中のすべての出現を未使用バイトコードに置き換えてから E を適用して、圧縮率が最良の隣接バイト対を置き換える隣接バイト対として選ぶ (直ぐ後及び3.3で L というパラメータを用いた拡張をする)。

その選んだ隣接バイト対について未使用バイトコードに置き換えた内容を新たに入力と考える。

置き換えられる未使用バイトコードが残っており、圧縮率が改善される間、以上を繰り返す。

I' は置き換えを逐次的に適用してでき上がると考えてもよいが、 I に I' にある置き換えをすべて適用した内容と等しい。 I' は、二分木による置き換えるの辞書と捉えることもできる。

後ほど、 I' と I'' の具体的なデータ形式について述べる。

本報告の方法では圧縮率の測定が必要になるが、ひとつの圧縮率のデータを測定するのに、いくつの置き換えを行っておくかという数を L と表すことにする (試すときに L 回置き換えできない場合はできる回数までとする)。5で L と K をパラメータとして圧縮率改善の実験を行っている。

効果のある最適な置き換えを探す方法はまだ分かっていない。本報告の実験では、未使用のバイトコードを使うごとに、最大 K^L 種類の E の試用の中から選定した。その制御は、 P のアルゴリズムの一部として、 P から機械的に行った。

I' の情報源としての性質は I'' とは異なることから、提案手法で E に直接適用するのは I'' のみとする。そのときの E の出力を $C_{I''}$ と表す。

ここで、混乱を避けるために断っておくと、導入の分かり易さのために用いてきた C_I 、 I^* 、 C_{I^*} という記号を、以下では用いないことにする。

3.2 提案手法の構成

以上をまとめると、符号化の構成は

入力は I , $I \rightarrow P \rightarrow I' + I''$, $I'' \rightarrow E \rightarrow C_{I''}$, 出力は $I' + C_{I''}$

↓
E

であり、復号の構成は

入力は $I'+C_{I''}$, $C_{I''} \rightarrow E^{-1} \rightarrow I''$,
 $I'+I'' \rightarrow P^{-1} \rightarrow I$, 出力は I

である。なお、符号化において P は E をいく度も使用するが、復号において P^{-1} からは E や E^{-1} を使用する必要はない。復号では、 I'' から、 I' を用いてワンプラスで I を得ることができる。

符号化には手間と時間をかけるが、復号はバイトデータを一度走査する程度のワンプラスの置き換えが増えるだけで済む。

P と連携したこのような E の使い方を、以下では簡単に $P \bullet E$ と記述する。

$P \bullet E$ を符号化として、 $(P \bullet E)^{-1}$ を復号として、それぞれひとまとまりとして考えると、通常の符号化、復号と同じ構造になる。

つまり、符号化は

入力は I , $I \rightarrow P \bullet E \rightarrow I'+C_{I''}$, 出力は $I'+C_{I''}$

であり、復号は

入力は $I'+C_{I''}$, $I'+C_{I''} \rightarrow (P \bullet E)^{-1} \rightarrow I$, 出力は I である。

3.3 前処理の具体的試行内容

実験で用いた前処理の試行内容を、具体例を用いて示す。圧縮対象の文字列の例として

a b a b c d a b c d ...

を用いる。これについて、探し当てた未使用のバイトコードの順序数のひとつを、仮に 129 とする。そのバイトデータのことを $\langle 129 \rangle$ と表す。2進数で表すなら 10000001 というデータである。他の順序数でも、個数の場合 (3.4 で出てくる) でも、同様に表記する。更に次に探し当てた未使用のバイトコードの順序数を、仮に 130 とする。

P の処理内容は、「隣接バイト対の中で圧縮率の面で効果的なものを未使用のバイトコードに置き換えておくこと」である。

置き換える隣接バイト対の出現頻度が高いほど、 I'' を短くすることになる。 I'' を短くすることが全体の圧縮率の改善と同義ではないが、本報告では、第ゼロ近似として、置き換えの候補として出現頻度の高いものから順に試している。

まず、 $L = 1$ の場合を述べ、後で拡張する。

P が置き換え内容を選ぶ処理の流れは、以下のとおりである。

- (1) 圧縮対象について隣接するバイト対すべてについての出現頻度を求め、頻度が高い隣接バイト対から K 種類を順に試しに置き換えをして、それぞれ E で圧縮して圧縮率を求める。
- (2) 試行した中で最良の結果を与えた組み合わせ

(隣接バイト対)を採用し、圧縮対象の中のその隣接バイト対のすべての出現において未使用のバイトコードに置き換える。

- (3) こうしてできたデータについて、新たに圧縮対象と捉えて、以上を繰り返す。未使用のバイトコードを使い切るか、改善しなくなったら繰り返しを終える。

$L = 1$, $K = 2$ の場合で、(1) と (2) を例示する。

圧縮対象において、頻度の高い方から $K = 2$ 種類の隣接バイト対として $a b$ と $b c$ が選ばれたとする。

まず、圧縮対象の中の $a b$ をすべて未使用のバイトコード $\langle 129 \rangle$ に置き換えた $\langle 129 \rangle \langle 129 \rangle c d \langle 129 \rangle c d \dots$ について E を使って圧縮率を調べる。

次に、圧縮対象の中の $b c$ をすべて未使用のバイトコード $\langle 129 \rangle$ に置き換えた $a b a \langle 129 \rangle d a \langle 129 \rangle d \dots$ について E を使って圧縮率を調べる。

最大 $K = 2$ 通りの試行の中で最も圧縮率が改善できたものを採用する。仮に、 $a b$ がそうであったとすると、圧縮対象の中の $a b$ を $\langle 129 \rangle$ に置き換えた

$\langle 129 \rangle \langle 129 \rangle c d \langle 129 \rangle c d \dots$

を新たに圧縮対象と捉えて以上のような処理を繰り返す。

L が 2 以上の場合については、(1) と (2) において次のように拡張する (ゲームの次の一手を決めるために、 L 手先まで手を読むことと類似している。探索は、枝が最大 K 本のゲーム木と同様である)。

隣接バイト対の置き換えを最大 K 種類試す際のそれぞれについて、更に隣接バイト対の置き換えを最大 K 種類試すということを、 L 重に行い、圧縮対象に L 重 (L 種類) の置き換えをした内容のそれぞれについて E を適用して圧縮率を調べる。新しい圧縮対象を作るために未使用のバイトコードと置き換えるのは、最も圧縮率を改善した L 重の置き換えの一重目の隣接バイト対である。

最大で K 通り試すことを L 重に行うため、 E を適用して圧縮率を調べる回数は、ひとつの未使用バイトコードあたり最大で K^L 通りになる。

これは一種の探索問題であり、局所最適解に陥る可能性はあるが、他によい方法を得ていない。

$L = 2$, $K = 2$ の場合で例示する。

圧縮対象について、頻度の高い方から $K = 2$ 種類の隣接バイト対として $a b$ と $b c$ が選ばれたとする。

まず、仮に圧縮対象のすべての $a b$ を $\langle 129 \rangle$ に置き換えた

$\langle 129 \rangle \langle 129 \rangle c d \langle 129 \rangle c d \dots$

について、頻度の高い方から $K = 2$ 種類の隣接バイト対として、仮に $\langle 129 \rangle c$ と $c d$ が選ばれたとする。そ

れぞれで $\langle 130 \rangle$ に置き換えをした $\langle 129 \rangle \langle 130 \rangle d \langle 130 \rangle d \dots$ と $\langle 129 \rangle \langle 129 \rangle \langle 130 \rangle \langle 129 \rangle \langle 130 \rangle \dots$ について E を使って圧縮率を調べる。

次に、仮に圧縮対象のすべての bc を $\langle 129 \rangle$ に置き換えた

$$a b a \langle 129 \rangle d a \langle 129 \rangle d \dots$$

について、頻度の高い方から $K = 2$ 種類の隣接バイト対として、仮に $a \langle 129 \rangle$ と $\langle 129 \rangle d$ が選ばれたとする。それぞれで $\langle 130 \rangle$ に置き換えをした $a b \langle 130 \rangle d \langle 130 \rangle d \dots$ と $a b a \langle 130 \rangle a \langle 130 \rangle \dots$ について E を使って圧縮率を調べる。

以上の調べた4つの中で、仮に3番目の圧縮率が良かったとすると、 bc を $\langle 129 \rangle$ に置き換えた

$$a b a \langle 129 \rangle d a \langle 129 \rangle d \dots$$

を新たに圧縮対象と捉えて以上のような処理を繰り返す。

3.4 置き換え情報 I' の表現

ここでは I' の表現形式について述べる。

I' として、まず、何種類置き換えをするか（仮に k 種類とする）をバイトコード $\langle k \rangle$ で表す。

もし置き換えをしないなら I' は

$$\langle 0 \rangle$$

の1バイトのみである。このように $I'+C_{I'}$ が1バイトは長くなる場合があるが、提案手法に効果がない場合でも1バイトの増加で済む。効果があれば、 I' がある程度の長さになっても、 $I'+C_{I'}$ 全体は短くなる。

置き換えるべきバイトコードの空きがない場合もこの表現を用いる。そうすると、復号側で置き換えるべきバイトコードの空きがない場合を扱うための処理は特別必要なくなる。

置き換えをする場合には、 $k \geq 1$ であり、 $\langle k \rangle$ の後ろに、 k 組分の「隣接バイト対と、置き換えをする空きのバイトコードの情報」を置く。それを少しでも短く表現するために、置き換えが何種類あるかによって、表現方法を2種類使い分ける。

置き換えが32種類以下の場合、隣接バイト対の2つのバイトコードと置き換えをする空きのバイトコードの計3バイトをひと組として必要なだけ並べる。

例えば、 ab を順序数129のバイトに、 $\langle 129 \rangle c$ を順序数130のバイトに置き換えたのであれば、 I' は

$$\langle 2 \rangle \underbrace{a b \langle 129 \rangle}_{32 \text{ バイト}} \underbrace{\langle 129 \rangle c \langle 130 \rangle}_{40 \times 2 \text{ バイト}}$$

の7バイトになる。ここで、波かっこは、分かりやすさのために付記したものであり、実際の情報には必要なく含まれない。以下でも同様である。

置き換えが32種類以上の場合、置き換えをする空きのバイトコードを、256ビット（すなわち、32バイト）のビットマップで表す。その後ろに、置き換えをする隣接バイト対すべてについて、バイト対をそのままバイトコードを2つ並べて表す。

例えば、先述の例を含め40種類を置き換えたのであれば、 I' は

$$\langle 40 \rangle \underbrace{\langle \cdot \rangle \langle \cdot \rangle \dots \langle \cdot \rangle}_{32 \text{ バイト}} a b \langle 129 \rangle c \underbrace{\langle \cdot \rangle \langle \cdot \rangle \dots \langle \cdot \rangle}_{40 \times 2 \text{ バイト}}$$

の113バイトの内容になっている。ここで、 $\langle \cdot \rangle$ は、具体的な内容の明示をしなかった適当な内容のバイトデータを表している。

I' の長さは k から

$$\begin{cases} 1 + 3k & (0 \leq k \leq 32) \\ 33 + 2k & (33 \leq k \leq 255) \end{cases}$$

と容易に知ることができる。

この報告では、使用していないバイトコードを探す必要があるものの、実際に探せばそのときに使用していないバイトコードは容易に分かることから、文字コードの種類が何であるか（例えば、JIS 8 単位コードかどうかというようなこと）が分かっている必要はない。

置き換えをするごとに I' が少しずつ長くなるが、使えるバイトコードの数は256を越えないため、 I' の長さは圧縮対象のデータ長に対して定数オーダーである。

3.5 提案手法に関する性質

提案手法の符号化は、もとにしたデータ圧縮プログラムが適応的であっても、全体としては一括型の符号化になる。

また、比較的効果的な隣接バイト対を探すことが必要なため、符号化の時間は桁違いに増すことがある。

提案手法の復号は元に戻す置き換えが必要になるが、これはバイトデータを置き換えながら一度走査する程度でよく、適応的にワンパスでも動作させられるため、パイプライン処理で既存符号の復号プログラムに渡すこともでき、その場合、既存符号の復号が適応的な処理である場合に、提案手法の全体としての復号の性質は適応的なままである。

符号化の時間は長いが復号は短いため、一度符号化したら幾度もまたは多数の箇所で使用するようなデータに向いている。

置き換えをしたことで I' はかく乱されているのではとの疑問も生じるが、記憶を持つ情報源の中で最も基本的なモデルである単純マルコフ情報源とい

う観点で見たときの平均情報量には影響がないことを確認している。データ圧縮の議論の中で述べると煩雑になるため、切り離して付録1から付録2で述べている。

4 記号対置き換えをする既存の取り組みとの比較

Gage⁴⁾のデータ圧縮プログラムは、隣接バイト対を頻度の高いものから順に空きのバイトコードに置き換えるというものである。これを更に他のバイトコード対応のデータ圧縮プログラムの入力とする、本報告のような利用はされていない。

本報告で試みた方法は、隣接バイト対を空きのバイトコードに置き換える点は同じであるが、目的のデータ圧縮プログラムEを使用して効果の高い隣接バイト対をある程度探索して選択する点が異なる。そして置き換えした内容を目的のデータ圧縮プログラムEで圧縮している。

文献5)の方法は、文字コードが既知のテキストファイルが対象で、日本語も含めて使われない62のバイトコードとその接続であるバイトコードに、2回以上現れた英数字列すべてを置き換えていた。方法は単純であるが、効果に関係なく置き換えを適用するため、中には大きく膨張してしまうファイルもあった。提案手法は、時間はかかるが効果を確かめながら処理する。最悪でも高々1バイト長くなるだけである。

5 実験と結果

実験には既存のデータ圧縮プログラムgzip (<http://www.gzip.org/>, version 1.3.5)、bzip2 (<http://www.bzip.org/>, version 1.0.5)及びctw (<http://www.ele.tue.nl/ctw/>, version 0.1)を用いた。圧縮率で優れている全く異なる方法の中から選んだ。

これらのアルゴリズムについて簡単に触れておく。gzipは、LZ77¹⁾とハフマン符号を用いている。bzip2は、ブロックソート法、MTF (Move-To-Front) 法及びハフマン符号を用いている²⁾。文脈木重み付け法 (Context-Tree Weighting)³⁾によるctwは独特の確率推定結果に算術符号を用いている。

圧縮性能は、ほとんどのファイルでgzip<bzip2<ctwであり、ctwが優る。圧縮の速さと一般的な利用の多さでは、gzip>bzip2>ctwであり、gzipが優る。

実験に用いた圧縮対象データは、Canterbury Corpusの中のテキストファイルである。ファイル名は実験結果の表中に記す。

実験結果を表1以降に示す。

表には各ファイルごとに3つのデータ圧縮プログラムについて、順に、既存プログラムのみの場合(Eのように表記。また、Eは必要なときに具体的なプログラム名に置き換えて表記。)の圧縮率Ratio(既存プログラムでの圧縮サイズをオリジナルのサイズで割って得ており、全く圧縮しない場合が100%)、提案手法(簡単にP•Eと表記)による圧縮率の改善 Δ Ratio(改善したバイト数とオリジナルのサイズとの比で、縮めた場合の符号は負)、圧縮率の改善 δ Ratio(改善したバイト数と使用した既存プログラムでの圧縮サイズとの比で、縮めた場合の符号は負)、改善したバイト数(縮めた場合の符号は負)、ひとつのファイルの圧縮中にPがEを呼び出した回数を記している。

効果がない場合には置き換えをしないため、どのデータ圧縮プログラムと併用しても1バイトよりも長くなることはない。

$L=1$ で、 K を1、2、3、10、100と変えた場合の実験結果を表1~表5に示す。表2~表5では、 K が増えたときにひとつ前の表よりも悪いデータに下線(カラー表示なら赤に見える彩色も)を施している。

K の増加に対して改善の効果は必ずしも単調増加ではないが、 K を増やすと多くの場合に改善の度合いが高まっている。中には、 δ Ratioで評価して、更に数%効果が出ている場合がある。

既存プログラムの使用回数は K を大きくするのに伴って増えている。

$L=1$ で $K=100$ とした場合と K^L が同じ値である、 $L=2$ で K が10の場合の実験結果を、表6に示す。これは、2種類の置き換え毎に圧縮率を評価したものである。表6では、表5よりも悪いデータに下線(カラー表示なら赤に見える彩色も)を施している。

L を増やした結果の優劣は一定していないが、 $L=1$ で $K=100$ の場合と比べて、良くなる場合が多い。言い換えると、制御がより単純な $L=1$ のアルゴリズムの方で、ある程度の圧縮率の改善が見込める。

6 おわりに

既存のデータ圧縮プログラムの適用前に圧縮対象に加工を加えて、全体としての符号化性能を上げられる場合のあることを示した。中には、優れた既存の圧縮手法の圧縮結果を、更に数%縮めている場合がある。

符号理論においては符号器をうまく構成して総合的に通信路をよりうまく使用することを探求してい

表 1 実験結果 ($L = 1, K = 1$)

File name	Size [Bytes]	gzip					bzip					ctw				
		P●Ratio	ΔRatio	δRatio	ΔSize	Ex. gzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. bzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. ctw
		[%]	[%]	[%]	[Bytes]	[Times]	[%]	[%]	[%]	[Bytes]	[Times]	[%]	[%]	[%]	[Bytes]	[Times]
alice29.txt	152089	35.79	-0.82	-2.28	-1240	16	28.41	0.00	0.00	1	2	25.94	0.00	0.00	1	2
asyoulik.txt	125179	39.10	0.00	0.00	1	2	31.61	0.00	0.00	1	2	29.03	0.00	0.00	1	2
cp.html	24603	32.52	0.00	0.01	1	2	30.99	0.00	0.01	1	2	28.84	0.00	0.01	1	2
fields.c	11150	28.19	-0.60	-2.13	-67	18	27.26	-0.15	-0.56	-17	8	24.88	-0.05	-0.22	-6	4
grammar.lsp	3721	33.40	-0.38	-1.13	-14	10	34.48	-0.11	-0.31	-4	4	29.80	-0.05	-0.18	-2	4
lcet10.txt	426754	33.95	-0.04	-0.12	-174	4	25.24	-0.03	-0.13	-138	4	22.90	-0.00	-0.01	-12	4
plrabn12.txt	481861	40.51	0.00	0.00	1	2	30.21	0.00	0.00	1	2	27.32	0.00	0.00	1	2
xargs.1	4227	41.57	0.02	0.06	1	2	41.68	0.02	0.06	1	2	37.02	0.02	0.06	1	2

“Size”:Original size,

“P”:Proposed preprocessor,

“Ratio”:(Compressed size by E)/(Original size),

“E”:Existing data compression program (gzip, bzip or ctw)

“ΔRatio”:(Compressed size by P●E)/(Original size) - (Compressed size by E)/(Original size)
 =(ΔSize)/(Original size),

“δRatio”:(Compressed size by P●E)/(Compressed size by E) - (Compressed size by E)/(Compressed size by E)
 =(ΔSize)/(Compressed size by E),

“ΔSize”:(Compressed size by P●E) - (Compressed size by E),

“Ex. E” :Execution count of E by P.

表 2 実験結果 ($L = 1, K = 2$)

File name	Size [Bytes]	gzip					bzip					ctw				
		P●Ratio	ΔRatio	δRatio	ΔSize	Ex. gzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. bzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. ctw
		[%]	[%]	[%]	[Bytes]	[Times]	[%]	[%]	[%]	[Bytes]	[Times]	[%]	[%]	[%]	[Bytes]	[Times]
alice29.txt	152089	35.79	-1.09	-3.06	-1663	24	28.41	-0.27	-0.96	-416	18	25.94	-0.04	-0.16	-65	6
asyoulik.txt	125179	39.10	-0.49	-1.26	-616	9	31.61	-0.08	-0.27	-105	6	29.03	0.00	0.00	1	3
cp.html	24603	32.52	0.00	0.01	1	3	30.99	0.00	0.01	1	3	28.84	0.00	0.01	1	3
fields.c	11150	28.19	-0.59	-2.10	-66	27	27.26	-0.15	-0.56	-17	12	24.88	-0.05	-0.22	-6	6
grammar.lsp	3721	33.40	-0.24	-0.72	-9	9	34.48	-0.11	-0.31	-4	6	29.80	-0.05	-0.18	-2	6
lcet10.txt	426754	33.95	-1.37	-4.04	-5856	45	25.24	-0.05	-0.19	-207	9	22.90	-0.00	-0.01	-12	6
plrabn12.txt	481861	40.51	-0.85	-2.10	-4098	15	30.21	-0.01	-0.03	-46	6	27.32	-0.01	-0.04	-51	6
xargs.1	4227	41.57	-0.21	-0.51	-9	12	41.68	0.02	0.06	1	3	37.02	0.02	0.06	1	3

表 3 実験結果 ($L = 1, K = 3$)

File name	Size [Bytes]	gzip					bzip					ctw				
		P●Ratio	ΔRatio	δRatio	ΔSize	Ex. gzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. bzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. ctw
		[%]	[%]	[%]	[Bytes]	[Times]	[%]	[%]	[%]	[Bytes]	[Times]	[%]	[%]	[%]	[Bytes]	[Times]
alice29.txt	152089	35.79	-1.10	-3.07	-1670	32	28.41	-0.25	-0.86	-373	20	25.94	-0.10	-0.40	-156	16
asyoulik.txt	125179	39.10	-0.75	-1.92	-938	28	31.61	-0.08	-0.27	-105	8	29.03	0.00	0.00	1	4
cp.html	24603	32.52	-0.17	-0.54	-43	20	30.99	0.00	0.01	1	4	28.84	0.00	0.01	1	4
fields.c	11150	28.19	-0.63	-2.23	-70	40	27.26	-0.15	-0.56	-17	16	24.88	-0.06	-0.25	-7	12
grammar.lsp	3721	33.40	-0.24	-0.72	-9	12	34.48	-0.11	-0.31	-4	8	29.80	-0.08	-0.27	-3	8
lcet10.txt	426754	33.95	-1.37	-4.04	-5853	60	25.24	-0.09	-0.37	-401	20	22.90	-0.01	-0.04	-41	12
plrabn12.txt	481861	40.51	-0.88	-2.17	-4233	24	30.21	-0.11	-0.36	-530	24	27.32	-0.02	-0.08	-99	12
xargs.1	4227	41.57	-0.21	-0.51	-9	16	41.68	0.00	0.00	0	8	37.02	0.02	0.06	1	4

表 4 実験結果 ($L = 1, K = 10$)

File name	Size [Bytes]	gzip					bzip					ctw				
		P●Ratio	ΔRatio	δRatio	ΔSize	Ex. gzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. bzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. ctw
alice29.txt	152089	35.79	-1.77	-4.94	-2690	198	28.41	-0.26	-0.91	-393	55	25.94	-0.10	-0.40	-156	44
asyoulik.txt	125179	39.10	-0.93	-2.38	-1163	110	31.61	-0.08	-0.27	-105	22	29.03	-0.00	-0.00	-1	22
cp.html	24603	32.52	-1.54	-4.75	-380	616	30.99	-0.07	-0.21	-16	33	28.84	-0.10	-0.34	-24	22
fields.c	11150	28.19	-0.90	-3.18	-100	154	27.26	-0.17	-0.63	-19	66	24.88	-0.13	-0.54	-15	55
grammar.lsp	3721	33.40	-0.27	-0.80	-10	44	34.48	-0.21	-0.62	-8	33	29.80	-0.08	-0.27	-3	22
lcet10.txt	426754	33.95	-2.37	-6.97	-10103	363	25.24	-0.24	-0.94	-1009	44	22.90	-0.01	-0.04	-41	33
plrabn12.txt	481861	40.51	-2.07	-5.11	-9977	352	30.21	-0.26	-0.85	-1234	110	27.32	-0.04	-0.14	-186	66
xargs.1	4227	41.57	-0.28	-0.68	-12	44	41.68	-0.09	-0.23	-4	22	37.02	0.02	0.06	1	11

表 5 実験結果 ($L = 1, K = 100$)

File name	Size [Bytes]	gzip					bzip					ctw				
		P●Ratio	ΔRatio	δRatio	ΔSize	Ex. gzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. bzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. ctw
alice29.txt	152089	35.79	-3.69	-10.31	-5610	14039	28.41	<u>-0.20</u>	<u>-0.71</u>	<u>-308</u>	404	25.94	-0.20	-0.77	-305	1515
asyoulik.txt	125179	39.10	-2.94	-7.52	-3681	14241	31.61	-0.22	-0.68	-270	505	29.03	-0.02	-0.06	-20	505
cp.html	24603	32.52	-1.85	-5.67	-454	8383	30.99	-0.10	-0.33	-25	404	28.84	-0.51	-1.76	-125	1616
fields.c	11150	28.19	-1.07	-3.79	-119	1919	27.26	-0.22	-0.79	-24	404	24.88	-0.39	-1.59	-44	1111
grammar.lsp	3721	33.40	-0.27	-0.80	-10	306	34.48	-0.21	-0.62	-8	157	29.80	-0.08	-0.27	-3	155
lcet10.txt	426754	33.95	-3.92	-11.55	-16734	17373	25.24	-0.33	-1.32	-1424	1414	22.90	-0.15	-0.64	-622	1212
plrabn12.txt	481861	40.51	-3.44	-8.49	-16569	13231	30.21	-0.36	-1.21	-1756	909	27.32	-0.07	-0.27	-354	1212
xargs.1	4227	41.57	-0.45	-1.08	-19	707	41.68	-0.14	-0.34	-6	202	37.02	0.02	0.06	1	101

表 6 実験結果 ($L = 2, K = 10$)

File name	Size [Bytes]	gzip					bzip					ctw				
		P●Ratio	ΔRatio	δRatio	ΔSize	Ex. gzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. bzip	P●Ratio	ΔRatio	δRatio	ΔSize	Ex. ctw
alice29.txt	152089	35.79	<u>-2.10</u>	<u>-5.86</u>	<u>-3192</u>	2886	28.41	-0.33	-1.15	-498	666	25.94	<u>-0.10</u>	<u>-0.40</u>	<u>-156</u>	444
asyoulik.txt	125179	39.10	<u>-1.02</u>	<u>-2.60</u>	<u>-1272</u>	1554	31.61	<u>-0.16</u>	<u>-0.50</u>	<u>-196</u>	555	29.03	<u>-0.00</u>	<u>-0.00</u>	<u>-1</u>	222
cp.html	24603	32.52	-1.88	-5.79	-463	7881	30.99	-0.15	-0.49	-37	555	28.84	<u>-0.10</u>	<u>-0.34</u>	<u>-24</u>	222
fields.c	11150	28.19	<u>-0.99</u>	<u>-3.50</u>	<u>-110</u>	2109	27.26	<u>-0.17</u>	<u>-0.63</u>	<u>-19</u>	666	24.88	<u>-0.17</u>	<u>-0.68</u>	<u>-19</u>	555
grammar.lsp	3721	33.40	-0.32	-0.97	-12	555	34.48	-0.21	-0.62	-8	333	29.80	-0.08	-0.27	-3	222
lcet10.txt	426754	33.95	<u>-2.41</u>	<u>-7.11</u>	<u>-10300</u>	4440	25.24	-0.40	-1.59	-1715	1998	22.90	<u>-0.07</u>	<u>-0.32</u>	<u>-309</u>	1776
plrabn12.txt	481861	40.51	<u>-2.79</u>	<u>-6.90</u>	<u>-13461</u>	10656	30.21	<u>-0.27</u>	<u>-0.88</u>	<u>-1285</u>	888	27.32	<u>-0.05</u>	<u>-0.18</u>	<u>-238</u>	888
xargs.1	4227	41.57	<u>-0.31</u>	<u>-0.74</u>	<u>-13</u>	888	41.68	<u>-0.09</u>	<u>-0.23</u>	<u>-4</u>	333	37.02	0.02	0.06	1	111

る。それは通信路の性質の理解や定式化が進んでいることから可能になっている。本研究では、既存のデータ圧縮プログラムが符号理論の通信路に相当し、前処理が符号器に相当する。前処理と関連付けたデータ圧縮プログラムの性質の理解や定式化が本質的な理論的課題といえる。

本報告は2つの発表(6)と7)をもとにまとめたものである。

謝辞

本研究についてご議論頂いた加治佐清光先生(当時、鹿児島工業高等専門学校情報工学科)に感謝いたします。

参考文献

- 1) J.Ziv and A.Lempel, "A Universal Algorithm for Sequential Data Compression," IEEE Trans. Inform. Theory, vol.IT- 23, no.3, pp.337-343, May 1977.
- 2) M.Burrows and D.J.Wheeler, "A Block-Sorting Lossless Data Compression Algorithm," SRC Research Report 124, Digital Systems Research Center, Palo Alto, May 1994.
- 3) F.M.J.Willems, Y.M.Shtarkov and T.J.Tjalkens, "The Context Tree Weighting Method: Basic Properties," IEEE Trans. Inform. Theory, vol.41, no.3, pp.653-664, May 1995.
- 4) P.Gage, "A new algorithm for data compression," The C Users Journal, R & D Publications, Inc. Lawrence, KS, USA, vol.12, no.2, pp.23-38, Feb. 1994.
- 5) 中村博文, 村島定行, "英数字列をひとまとまりとして扱うデータ圧縮について," 信学論A, Vol. J80-A, N0.9, pp.1559-1563, Sept. 1997.
- 6) 中村博文, 加治佐清光, 澁田孝康, "データ圧縮の効率を高める前処理の可能性について," 情報理論とその応用シンポジウム予稿集31巻, 論文番号6.4.4, pp 698-703, Oct. 2008.
- 7) 中村博文, 加治佐清光, 澁田孝康, "テキストデータの圧縮効率を高める前処理の可能性について," 電子情報通信学会情報理論研究会技術報告IT2009巻123号, pp.337-342, Mar. 2010.

付録1 置き換え内容の表現に必要な情報量の変化

一般化して記号対の置き換えの影響を確認する。

以下では、圧縮対象の1記号当たりの、ある前提とした情報源モデルにおいての、平均情報量のこと

を単にエントロピーと呼び H と表す。置き換えると内容が変わるが、その場合でも、置き換え前の元の圧縮対象の1記号当たりで、表現に必要な平均情報量やエントロピーと呼ぶ。

まず、全体としては可逆であるため、置き換えによって情報が失われることはない。すなわち、置き換えた内容のエントロピーが小さくなることはない。

しかし、置き換えた内容は加工されているため、置き換えた内容を表すのに必要なビット数が、情報源モデルによっては、元の圧縮対象の1記号当たりで考えてエントロピーより多い記述量が必要になるのではないかという疑問が生じる。

加工の活用の例として、分野は異なるが、誤り訂正符号の利用ではインタリーブによってバースト的な誤りへの耐性を上げられる。しかし、データ圧縮では勝手な記号のシャッフルは、通常は効果的な圧縮の邪魔になる。

シャッフルよりは簡単な加工である置き換えについて見ていく。置き換えた結果の表現に必要な平均情報量を、但し、置き換え前の元の圧縮対象の1記号当たりでの平均情報量を、 G と表す。

前提とする情報源モデルによっては $G > H$ ではないかという疑問が生じる。圧縮対象のそれぞれのファイルの真の情報源モデルは分かるものではないが、本報告では、基本的な定式化として、単純マルコフ情報源として捉えた場合について付録2で検証する。

結論から述べると、 $G = H$ である。

付録2 単純マルコフ情報源として置き換え内容を捉える場合

単純マルコフ情報源を通常と同じく状態遷移で捉えることにする。

使用記号の集合を \mathcal{A} 、状態の集合を \mathcal{S} と表す。状態 $x(x \in \mathcal{S})$ を取る確率を P_x 、状態 $x(x \in \mathcal{S})$ から記号 $i(i \in \mathcal{A})$ で遷移する確率を p_{xi} のように表す。状態 $x(x \in \mathcal{S})$ から記号 $i(i \in \mathcal{A})$ で遷移した状態を $[x, i]$ のように表すことがある。そこから更に遷移した状態はコマで区切って書き足して表す。例えば、 $[x, i]$ から $j(j \in \mathcal{A})$ で遷移した状態を $[x, i, j]$ のように表す。

状態 $y(y \in \mathcal{S})$ から遷移するときのエントロピーを H_y と表す。 $H_y(y \in \mathcal{S})$ は、

$$H_y = \sum_{j \in \mathcal{A}} -p_{yj} \log_2 p_{yj} \quad (1)$$

である。 \sum の直後の項に単項演算子がある場合に、カッコで囲むことは省略する。 $-p_{yj} \log_2 p_{yj}$ はエント

ロピーの構成要素である。

単純マルコフ情報源のエントロピー H は

$$H = \sum_{y \in \mathcal{S}} P_y H_y \quad (2)$$

である。ここで、

$$\begin{aligned} P_y &= \sum_{x \in \mathcal{S}} \sum_{i \in \mathcal{A}, [x,i]=y} P_x p_{xi} \\ &= \sum_{x \in \mathcal{S}} P_x \sum_{i \in \mathcal{A}, [x,i]=y} p_{xi} \end{aligned} \quad (3)$$

であることを用いて、

$$\begin{aligned} H &= \sum_{y \in \mathcal{S}} \sum_{x \in \mathcal{S}} P_x \sum_{i \in \mathcal{A}, [x,i]=y} p_{xi} H_y \\ &= \sum_{x \in \mathcal{S}} P_x \sum_{y \in \mathcal{S}} \sum_{i \in \mathcal{A}, [x,i]=y} p_{xi} H_y \\ &= \sum_{x \in \mathcal{S}} P_x \sum_{i \in \mathcal{A}} p_{xi} H_{[x,i]} \\ &= \sum_{x \in \mathcal{S}} \sum_{i \in \mathcal{A}} P_x p_{xi} H_{[x,i]} \end{aligned} \quad (4)$$

と表せる。

$P_x p_{xi} H_{[x,i]}$ は遷移全体の中での、 x から i で遷移する確率と、その遷移先のエントロピーをかけた値である。 $P_x p_{xi} H_{[x,i]}$ はエントロピーの構成要素である

付録 2.1 単純マルコフ情報源についての置き換え

今、 x が u 、 i が a 、 y が $[u, a]$ のときと対応する

$$\begin{array}{c} \downarrow \cdots \downarrow \\ (u) - a \rightarrow ([u, a]) \\ || \cdots | \\ j(\in \mathcal{A}) \\ \downarrow \cdots \downarrow \end{array}$$

のような連なりに、すなわち遷移に、注目する。状態はカッコの中に記している（以下同様）。ここで、 u からは $[u, a]$ への遷移のみ考えているが、 $[u, a]$ へは他の状態からの遷移もあり得る。

$[u, a]$ からはすべての $j(\in \mathcal{A})$ についての遷移がある。このことは、上記において、簡潔に

$$\begin{array}{c} || \cdots | \\ j(\in \mathcal{A}) \\ \downarrow \cdots \downarrow \end{array}$$

のように表記している。

$[u, a]$ からの b での遷移を分けると

$$\begin{array}{c} \downarrow \cdots \downarrow \\ (u) - a \rightarrow ([u, a]) - b \rightarrow \\ || \cdots | \\ j(\in \mathcal{A}, \neq b) \\ \downarrow \cdots \downarrow \end{array}$$

と表せる。単純マルコフ情報源の枠組み内で、 $[u, a]$ への u からの遷移の分だけを、新たな状態の表記（ここでは v ）を設けて分けることができる。すなわち、

$$\begin{array}{c} (u) - a \rightarrow (v) - b \rightarrow \\ || \cdots | \\ j(\in \mathcal{A}, \neq b) \\ \downarrow \cdots \downarrow \end{array}$$

のように捉える。以下では、 $j(\in \mathcal{A}, \neq b)$ の部分、または $j(\in \mathcal{A}, \neq a)$ の部分は一層簡単に「 $\downarrow \downarrow$ 」と表す。

u から v を選ぶことと、 v についての平均情報量は

$$\begin{aligned} h_{ua} &= P_u(-p_{ua} \log_2 p_{ua}) + P_u p_{ua} H_{[u,a]} \\ &= P_u(-p_{ua} \log_2 p_{ua}) + P_u p_{ua} H_v \quad (5) \\ &= P_u(-p_{ua} \log_2 p_{ua}) \end{aligned}$$

$$+ P_u p_{ua} \sum_{j \in \mathcal{A}} -p_{vj} \log_2 p_{vj} \quad (6)$$

である。

単純マルコフ情報源において、任意のバイトコードによる隣接記号対（そのひと組を ab とする）をひとつの未使用バイトコード r に置き換える場合を考える。

ab を r に置き換えても v からの遷移として考えられるすべての場合について表現に必要な平均情報量がエントロピー H と同じであることを示すには、 h_{ua} だけに関連した部分について置き換えても単純マルコフ情報源として見ているときの表現に必要な平均情報量 (g_{ua} と表す) が変わらないことを示せばよい。すなわち、 $g_{ua} = h_{ua}$ によって $G = H$ を言う。

以下に場合分けして示す。

以下では、確率（仮に、 p である場合）について $1-p$ を \bar{p} のように表す。

付録 2.2 a で始まる $a \neq b$ 、 $u \neq v$ 、 $[v, b] \neq v$ を含む場合

まず、最も簡単な、 $a \neq b$ 、 $u \neq v$ 、 $[v, b] \neq v$ を含む場合について考える。

u から v を通過する部分だけに注目すると

$$\begin{array}{c} (u) - a \rightarrow (v) - b \rightarrow \\ \downarrow \downarrow \end{array}$$

のように連なっている。なお、ここで、 v へは u からの遷移のみ考えるが、 v からはすべての $j(\in \mathcal{A})$ の遷

移を考慮する。これは、

$$\begin{array}{c} (u) - a \rightarrow (v') - b \rightarrow \\ \quad \downarrow \quad \quad \quad \quad \downarrow \\ \quad \quad a \rightarrow (v'') \\ \quad \quad \quad \quad \quad \quad \quad \downarrow \downarrow \end{array}$$

のように表せる。更に、ここでの上側の経路の遷移では、 a の後ろでは b が確率1で接続するので、記号対 ab を \mathcal{A} の中でそれまでに使われていなかった r と書き変えて、(但し、 r を \mathcal{A} に加えたものを新しく \mathcal{A} と表すのは証明が一旦済んでからとする)

$$\begin{array}{c} (u) \text{ --- } r \rightarrow \\ \quad \downarrow \quad \quad \quad \quad \downarrow \\ \quad \quad a \rightarrow (v'') \\ \quad \quad \quad \quad \quad \quad \quad \downarrow \downarrow \end{array}$$

のように書き換えられる。(v'' からの $j \in \mathcal{A}, j \neq b$ での遷移確率は $\frac{p_{vj}}{p_{vb}}$ である。)

この部分を表現するのに必要な平均情報量を g_{ua} とする。以下、 \log_2 の直後の積にカッコは省略する。 g_{ua} は、

$$\begin{aligned} g_{ua} &= P_u p_{ua} \times 0 \\ &+ P_u (-p_{ua} p_{vb} \log_2 p_{ua} p_{vb}) \\ &+ P_u (-p_{ua} \overline{p_{vb}} \log_2 p_{ua} \overline{p_{vb}}) \\ &+ P_u p_{ua} \overline{p_{vb}} \sum_{j \in \mathcal{A}, j \neq b} -\frac{p_{vj}}{p_{vb}} \log_2 \frac{p_{vj}}{\overline{p_{vb}}} \\ &= P_u (-p_{ua} \log_2 p_{ua}) \\ &+ P_u p_{ua} (-p_{vb} \log_2 p_{vb} - \overline{p_{vb}} \log_2 \overline{p_{vb}}) \\ &+ P_u p_{ua} \left(\sum_{j \in \mathcal{A}, j \neq b} -p_{vj} \log_2 p_{vj} \right. \\ &\quad \left. + \left(\sum_{j \in \mathcal{A}, j \neq b} p_{vj} \right) (\log_2 \overline{p_{vb}}) \right) \\ &= P_u (-p_{ua} \log_2 p_{ua}) \\ &\quad + P_u p_{ua} \sum_{j \in \mathcal{A}} -p_{vj} \log_2 p_{vj} \\ &= P_u (-p_{ua} \log_2 p_{ua}) + P_u p_{ua} H_v \end{aligned} \quad (7)$$

となり、 h_{ua} と等しい。よって、 $G = H$ である。

付録2.3 a で始まる $a \neq b$ 、 $u \neq v$ 、 $[v, b] = v$ を含む場合

次に、 $a \neq b$ 、 $u \neq v$ 、 $[v, b] = v$ を含む場合について考える。

u から v を通過する部分だけに注目すると

$$\begin{array}{c} (u) - a \rightarrow (\overline{v}) - b \rightarrow \\ \quad \quad \quad \quad \quad \quad \quad \downarrow \downarrow \end{array}$$

のように連なっている。なおここで、 v からはすべての $j \in \mathcal{A}$ の遷移を考慮する。 v へは u からの遷移と v からの遷移がある。ここで、エントロピーの構成要素の内、状態 v に関する量は

$$P_v H_v = P_v \sum_{j \in \mathcal{A}} -p_{vj} \log_2 p_{vj}$$

である。先述の連なりは、 v から自身への遷移を1回分だけほどいて2つの状態 w' と w'' を用いて表し直すと

$$\begin{array}{c} (u) - a \rightarrow (w') - b \rightarrow (\overline{w''}) - b \rightarrow \\ \quad \quad \quad \quad \quad \quad \quad \downarrow \downarrow \quad \quad \quad \downarrow \downarrow \end{array}$$

と表すことができる。 $H_{w''} = H_v$ である。

ここで、エントロピーの構成要素の内、状態 w' と w'' に関する量は

$$p_v \overline{p_{vb}} \sum_{j \in \mathcal{A}} -p_{vj} \log_2 p_{vj} + P_v p_{vb} H_v = P_v H_v \quad (8)$$

となり、記号対の置き換えによって変化はない。

ここでの u から w'' までの連なりは前節と同じであるので、 $[v, b] = v$ の場合でも、記号対 ab を r に書き変えても $H = G$ である。

付録2.4 a で始まる $a \neq b$ 、 $u = v$ 、 $[v, b] \neq v$ を含む場合

a による自身への遷移があり得る場合である。

紙面の都合で証明を略すが、 $H = G$ である。

付録2.5 a で始まる $a \neq b$ 、 $u = v$ 、 $[v, b] = v$ を含む場合

a と b による自身への遷移があり得る場合である。

紙面の都合で証明を略すが、 $H = G$ である。

付録2.6 a で始まる $a = b$ 、 $u \neq v$ 、 $[v, b] \neq v$ を含む場合

同じ記号での記号対の場合であり、更に自身への遷移がない場合である。

なお、ここでの議論には、 $[v, a] = u$ の場合も含まれる。

u から v を通過する部分だけに注目すると

$$\begin{array}{c} (u) - a \rightarrow (v) - a \rightarrow \\ \quad \quad \quad \quad \quad \quad \quad \downarrow \downarrow \end{array}$$

のように連なっている。なおここで、 v へは u からの遷移のみ考えるが、 v からはすべての $j \in \mathcal{A}$ の遷移を考慮する。これは、

$$(u) - a \rightarrow (v') - a \rightarrow$$

$$\quad \downarrow \quad \downarrow$$

$$\quad \quad a \rightarrow (v'')$$

$$\quad \quad \downarrow \downarrow$$

のように表せる。更に、これの上側の経路の遷移では、 a の後ろで a が確率1で接続するため、記号対 aa を \mathcal{A} の中でそれまでに使われていなかった r と書き変えて、(但し、 r を \mathcal{A} に加えたものを新しく \mathcal{A} と表すのは証明が一旦済んでからとする)

$$(u) - r - \rightarrow$$

$$\quad \downarrow \quad \downarrow$$

$$\quad \quad a \rightarrow (v'')$$

$$\quad \quad \downarrow \downarrow$$

のように書き換えられる。 $(v''$ からの $j(\in \mathcal{A}, \neq a)$ での遷移確率は $\frac{p_{vj}}{p_{va}}$ である。)

ここから先は $a \neq b$ の議論と同様にして、 aa を r に置き変えても表現に必要な平均情報量が変化しないことが言える。

付録 2.7 a で始まる $a = b$ 、 $u \neq v$ 、 $[v, b] = v$ を含む場合

同じ記号での記号対の場合で、更に2番目で自身への遷移があり得る場合である。

u から v を通過する部分だけに注目すると

$$(u) - a \rightarrow (v) - a - \rightarrow$$

$$\quad \downarrow \quad \downarrow$$

$$\quad \quad \downarrow \quad \downarrow$$

のように連なっている。なおここで、 v からはすべての $j(\in \mathcal{A})$ の遷移を考慮する。 v へは u からの遷移と v からの遷移がある。

P_v について、 $P_v = P_u p_{ua} + P_v p_{va}$ という関係が成り立ち、解くと

$$P_v = \frac{P_u p_{ua}}{p_{va}} \tag{9}$$

である。

よって、エントロピーの構成要素の内、 u から v を選ぶことと、 v に関する量は、 $P_u(-p_{ua} \log_2 p_{ua}) + P_v H_v$ である。

上記の連なりは、まず $a \neq b$ の場合に行ったように、 v から v への遷移を1回分だけほどいて2つの状態を用いて表し直すと

$$(u) - a \rightarrow (\cdot) - a \rightarrow (\cdot) - a - \rightarrow$$

$$\quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

$$\quad \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

と表すことができる。名前を付けない状態については、中点を記している。ここでは、 $[u, a]$ と $[u, a, a]$ がそれに当たる。 $[u, a]$ を取る確率 $P_u p_{ua}$ は P_v を用いて表すと式(9)より $P_v \overline{p_{va}}$ であり、 $[u, a, a]$ を取る確

率は P_v から $[u, a]$ を取る確率を引いて $P_v p_{va}$ である。また、 $H_{[u, a]} = H_{[u, a, a]} = H_v$ である。

上記の連なりの、後ろの a での繰り返しの部分も考慮する必要がある。一旦、 a での繰り返しの部分を、無限にほどいたと考えると、

$$(u) - a \rightarrow (\cdot) - a \rightarrow (\cdot) - a \rightarrow (\cdot) - a \rightarrow (\cdot) \dots$$

$$\quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

と表すことができる。 u の後ろの第 k 番目の状態が取る確率は P_v を用いて表すと、 $P_v p_{va}^{k-1} \overline{p_{va}}$ である。

これの2番目以降の偶数番目と、3番目以降の奇数番目とをまとめると、

$$(u) - a \rightarrow (\cdot) - a \rightarrow (\cdot) - a \rightarrow (\cdot) - a - \rightarrow$$

$$\quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

と表すことができる。

これの u の後ろの状態を取る確率は $P_v \overline{p_{va}}$ である。その後ろは、先の無限の状態の2番目の状態を取る確率 $P_v p_{va} \overline{p_{va}}$ に、 $1 + x + x^2 + \dots = \frac{1}{1-x}$ なる関係を用いて偶数番目をすべて足し合わせて、

$$P_v p_{va} \overline{p_{va}} \frac{1}{1-p_{va}^2} = \frac{P_v p_{va}}{1+p_{va}}$$

である。更にその後ろは、先の無限の状態の3番目の状態を取る確率をもとにして、

$$P_v p_{va}^2 \overline{p_{va}} \frac{1}{1-p_{va}^2} = \frac{P_v p_{va}^2}{1+p_{va}}$$

である。(ここでは確認として、 u の後ろのこれら3つの状態のどれかを取る確率を求めると、3者を足して P_v である。記号対の置き換えが他の状態の確率に影響を与えていないと言える。)

aa を r に書き変えると

$$(u) - r - \rightarrow (v'') - r - \rightarrow$$

$$\quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

$$\quad \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

のように書き換えられる。 v' 、 v'' 、 v''' はここで新たに定義した名前である。 $(v', v'', v'''$ からの $j(\in \mathcal{A}, \neq a)$ での遷移確率は $\frac{p_{vj}}{p_{va}}$ である。)

v' 、 v'' 、 v''' を取る確率はそれぞれ $P_v \overline{p_{va}}^2$ 、 $\frac{P_v p_{va}}{1+p_{va}}$ 、 $\frac{P_v p_{va}^2 \overline{p_{va}}}{1+p_{va}}$ である。

エントロピーの構成要素の内、 u から上記の後ろの部分を選ぶことと、後ろの部分の状態に関する量を簡単に h とおく。 h は

$$h = P_u p_{ua} \times 0$$

$$+ P_u (-p_{ua} p_{va} \log_2 p_{ua} p_{va})$$

$$\begin{aligned}
 & +P_u(-p_{ua}\overline{p_{va}}\log_2 p_{ua}\overline{p_{va}}) \\
 & +P_v\overline{p_{va}}^2 \sum_{j \in \mathcal{A}, j \neq a} -\frac{p_{vj}}{p_{va}} \log_2 \frac{p_{vj}}{p_{va}} \\
 & +\frac{P_v p_{va}}{1+p_{va}} \left(-p_{va}^2 \log_2 p_{va}^2 - p_{va}\overline{p_{va}} \log_2 p_{va}\overline{p_{va}} \right. \\
 & \qquad \qquad \qquad \left. - \sum_{j \in \mathcal{A}, j \neq a} -p_{vj} \log_2 p_{vj} \right) \\
 & +\frac{P_v \overline{p_{va}}^2}{1+p_{va}} \sum_{j \in \mathcal{A}, j \neq a} -\frac{p_{vj}}{p_{va}} \log_2 \frac{p_{vj}}{p_{va}} \\
 = & P_u(-p_{ua}p_{va} \log_2 p_{ua} - p_{ua}\overline{p_{va}} \log_2 p_{va} \\
 & \qquad - p_{ua}\overline{p_{va}} \log_2 p_{ua} - p_{ua}\overline{p_{va}} \log_2 \overline{p_{va}}) \\
 & +P_v\overline{p_{va}} \sum_{j \in \mathcal{A}, j \neq a} -p_{vj} \log_2 p_{vj} \\
 & -P_v\overline{p_{va}} \left(\sum_{j \in \mathcal{A}, j \neq a} -p_{vj} \right) \log_2 \overline{p_{va}} \\
 & +\frac{P_v p_{va}}{1+p_{va}} \left(-p_{va}^2 \log_2 p_{va} - p_{va}^2 \log_2 p_{va} \right. \\
 & \qquad \qquad \qquad \left. - p_{va}\overline{p_{va}} \log_2 p_{va} - p_{va}\overline{p_{va}} \log_2 \overline{p_{va}} \right. \\
 & \qquad \qquad \qquad \left. - \sum_{j \in \mathcal{A}, j \neq a} -p_{vj} \log_2 p_{vj} \right) \\
 & +\frac{P_v \overline{p_{va}}^2}{1+p_{va}} \left(\sum_{j \in \mathcal{A}, j \neq a} -p_{vj} \log_2 p_{vj} \right. \\
 & \qquad \qquad \qquad \left. - \sum_{j \in \mathcal{A}, j \neq a} -p_{vj} \log_2 \overline{p_{va}} \right) \\
 = & P_u(-p_{ua} \log_2 p_{ua}) \\
 & +P_v \left(\overline{p_{va}} H_v + \frac{p_{va}^2}{1+p_{va}} H_v + \frac{p_{va}}{1+p_{va}} H_v \right) \\
 = & P_u(-p_{ua} \log_2 p_{ua}) + P_v H_v \tag{10}
 \end{aligned}$$

となり、よって、 aa を r に置き換える前と比べて表現に必要な平均情報量は変わらない。

付録 2.8 a で始まる $u = v$ を含む場合

同じ記号での記号対の場合で、更に最初に自身への遷移があり得る場合である。

残っていた2つのケースについて、紙面の都合で証明を略すが、 $[v, b] \neq v$ でも、 $[v, b] = v$ でも、 $H = G$ である。

付録 3 m 重マルコフ情報源として置き換え内容を捉える場合

念のために明記しておく、 $m = 1$ の場合は単純

マルコフ情報源であり付録 1 から付録 2 で述べている。この節で言及するのは $m \geq 2$ の場合である。

ab を r に置き換えた内容の各場所で、過去の m 記号から次のひとつの確率を知ることができるため、証明自体は可能である。しかし、場合分けが多くなりすぎてすべてを列記できないため、証明における場合分けの方針だけ述べておく。

m 重マルコフ情報源として捉える場合、先の記号について、 a と r (すなわち、 ab) は区別が必要である。置き換え前に

$$\underbrace{\dots\dots\dots a}_{m \text{ 記号}}$$

であったものが、置き換えると

$$\underbrace{\dots\dots\dots a}_{m \text{ 記号}} (\neq \underbrace{\dots\dots\dots ab}_{m \text{ 記号}})$$

及び

$$\underbrace{\dots\dots\dots r}_{m \text{ 記号}}$$

になる。

これについて、 r は ab と続く確率で遷移し、 a は ab と続く確率を除いた確率で遷移する。

過去の m 記号に ab が含まれていた場合の置き換え後の m 記号の状態からの繊維というのは、実質的には元々の m 個より多い記号での状態から遷移することになる。また、それぞれの状態からの遷移確率は重複や欠けがなく量的に知ることができる。これは、例えば、過去の m 記号に ab が1対だけ含まれる場合なら、置き換え前に

$$\underbrace{\dots a b \dots a}_{m \text{ 記号}}$$

であったものが、置き換えると

$$\underbrace{\dots r \dots a}_{m \text{ 記号}} (\neq \underbrace{\dots r \dots ab}_{m \text{ 記号}})$$

及び

$$\underbrace{\dots r \dots r}_{m \text{ 記号}}$$

になる。

過去の m 記号に含まれる ab の対 (すなわち、 r) の数を2つ、3つ、... と最大 m まで増やして網羅して場合分けすればよい。

Cu₂SnS₃ 薄膜太陽電池における高効率化の検討

内村友宏¹・荒木秀明²・中村重之³・赤木洋二⁴

Investigation of high efficiency in Cu₂SnS₃ thin film solar cells

UCHIMURA Tomohiro¹, ARAKI Hideaki², NAKAMURA Shigeyuki³, and AKAKI Yoji⁴

(Accepted September 27, 2021)

Abstract In this study, we investigated the effect of H₂S concentration on Cu₂SnS₃ (CTS) thin films and the film thickness of CdS buffer layer to further improve conversion efficiency. CTS thin films were fabricated by sulfurization of Cu/Sn/glass stacked precursors deposited by a vacuum evaporation. After deposition, the precursor was crystallized by annealing in H₂S atmosphere for 20 minutes at 500 °C. The H₂S concentration was changed from 2 to 6%. From Hall effect measurement, the highest mobility and lowest carrier density were obtained at the H₂S concentration of 2%. For examination of CdS thickness, thin films were prepared by changing deposition time from 10 to 16 minutes using the chemical bath deposition method. The highest conversion efficiency of 1.16% was obtained at for 14 minutes. It is considered that the leakage current decreased due to the increase in film thickness.

Keywords [Thin film solar cells, Cu₂SnS₃, Vacuum evaporation, H₂S annealing]

1 序論

持続可能な低炭素社会の実現に向け、太陽光発電への期待はますます高まっており、2050年には現在の約10倍である数TW(テラワット)の年間導入量になると予測されている。しかし、現在主流である結晶シリコン太陽電池の高い生産コストや原料枯渇の可能性を考えると、この要求に応えることは難しい。そこで近年、化合物薄膜太陽電池が注目を集めている。化合物薄膜太陽電池は、薄膜化が可能な光吸収係数の大きな材料を利用するため省資源で、製造コストも安く、化合物の組み合わせによって原料コス

トを安くできるという特徴を持っている。現在製品化されているものとしては Cu(In,Ga)(S,Se)₂ (CIGS) や CdTe があるが、それらは希少金属の In、Ga、Se、Te、有毒元素の Cd、Te を使用しており、安定供給や環境負荷の面で不安がある。このような背景から本研究室では、将来の大規模導入を見据え、希少金属や有毒元素を使用しない Cu₂SnS₃ (CTS)半導体に注目している。この材料は、10⁴cm⁻¹以上の高い光吸収係数と 0.93eV のバンドギャップを有していることから¹⁾、2 接合タンデム型太陽電池のボトム層としての使用が期待されている。例えば、CTSをボトム層に、Cu₂ZnSnS₄(CZTS)をトップ層に用いた場合、

1 東北大学工学研究科 Graduate School of Engineering, Tohoku University

2 長岡工業高等専門学校物質工学科 Department of Materials Engineering, National Institute of technology(KOSEN), Nagaoka College

3 津山工業高等専門学校総合理工学科 Department of Integrated Science and Technology, National Institute of technology(KOSEN), Tsuyama College

4 都城工業高等専門学校電気情報工学科 Department of Electrical and Computer Engineering, National Institute of technology(KOSEN), Miyakonjo College

理論変換効率は 40%以上であり²⁾、高効率・低コスト・低環境負荷の太陽電池が実現可能となる。しかしながら、CTS 薄膜太陽電池の最高変換効率は 5.1%と低く³⁾、実用化のためには 10%以上変換効率を向上させる必要がある。この低い変換効率の要因としては、結晶粒径が小さいこと、蒸気圧の高い SnS の再蒸発などによるポイドの形成、表面荒さが大きいこと等が挙げられる⁴⁾。

そこで我々は、更なる膜の高品質化を狙うために、H₂S ガスを利用した熱処理を行うことで硫化及びアニールを行う、H₂S アニールに着目した。このアニールの特徴としては、抜けやすい S 原子を一定量供給できること、固体硫黄を使った硫化に比べて省資源であることが挙げられる。このような特徴を有していることから CIGS の製造プロセスにも使用されている⁵⁾。さらに、化学浴体積(CBD)法による CdS 薄膜の成膜条件は、CTS 薄膜の表面形態と密接に関係しており、CTS 薄膜の成膜条件にあった CdS 薄膜の成膜条件を検討する必要がある。

よって、本研究では、H₂S 濃度を変化させることで CTS 薄膜の高品質化を狙うとともに、CdS 薄膜の膜厚の最適化を行うことで、更なる高効率化を目指すことを研究目的とした。

2 実験

2.1 CTS 薄膜における H₂S 濃度の検討

基板温度を Cu、Sn とともに室温とし、真空蒸着法により Cu/Sn/glass プリカーサを成膜した。この時、蒸着時の仕込み組成比を Cu/Sn=1.75 とした。その後、作製したプリカーサ膜を N₂ と H₂S の混合ガス雰囲気中で 500°C、20 分間アニールを行い、CTS 薄膜を作製した。この時、混合ガス中の H₂S 濃度を 2、4、6%と変化させることで H₂S 濃度に対する CTS 薄膜の結晶の品質を検討した。

2.2 CdS 薄膜における膜厚の検討

CdS 薄膜における膜厚の検討を行うため、glass/Mo/Cu₂SnS₃/CdS/ZnO:Al/Al 構造の太陽電池を作製し、その影響を調べた。Mo 付き glass 基板に、電子ビーム蒸着により glass/Mo/Sn/Cu プリカーサ膜を成膜し、S 粉末と共に 570°C で 2 時間のアニールを行うことで CTS の成膜を行った。その後、CdS 薄膜は化学浴堆積(CBD)法、Al 添加 ZnO は RF スパッタリング法、Al 電極は真空蒸着法より作製を行った。ここで、CdS の成膜時間を 10、12、14、16 分とし、膜厚を変化させた。

2.3 評価方法

作製した薄膜の結晶構造をリガク SmartLab X線回折(XRD)装置、表面形態を日立 SU8020 走査型電子顕微鏡(SEM)、元素マッピング・組成分析を BRUKER QUANTAX FlatQUAD 5060 エネルギー分散型 X線分析装置(EDS)、電気的特性を Nanometrics HL5500 ホール効果測定装置にて評価を行った。さらに、デバイスの太陽電池特性を JASCO YQ-250BX ソーラーシミュレータ(AM1.5 100mW/cm²)、外部量子効率を分光計器 SML-250J 分光感度測定装置(波長範囲 : 300~1400nm)にて評価を行った。

3 実験結果と考察

3.1 CTS 薄膜における H₂S 濃度の検討

Fig. 1 に異なる H₂S 濃度で作製した CTS 薄膜の XRD パターンを示す。全ての H₂S 濃度で単斜晶構造 CTS に帰属される回折ピークと SnS 結晶に帰属される回折ピークが確認された。また、6%では、unknown ピークも確認された。Fig. 2 に H₂S 濃度を変化させた時の CTS と SnS の第一ピーク強度を示す。H₂S 濃度を増加させると、CTS ピーク強度は減少し、SnS ピーク強度は増加する結果となった。これは、H₂S 濃度増加によって、Sn と S の反応がさらに増加し、SnS の異相形成に繋がっていると考える。結果的に、H₂S 濃度 2%において最も SnS ピーク強度が減少し、単相に近い薄膜が得られた。

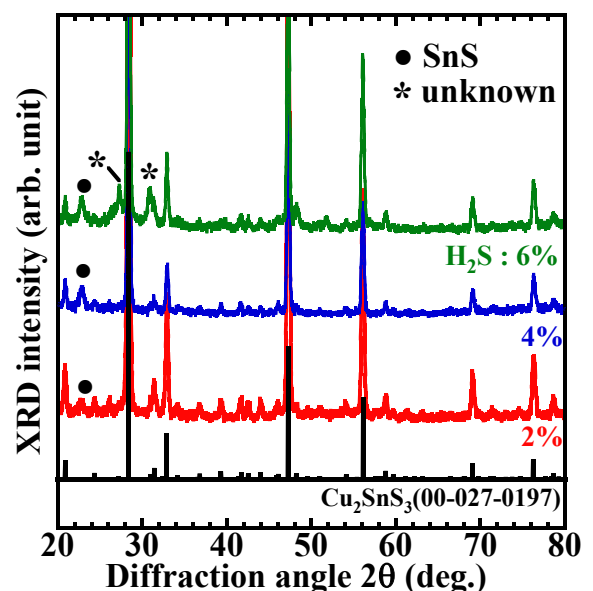


Fig. 1 XRD patterns of CTS thin films prepared at different H₂S concentrations

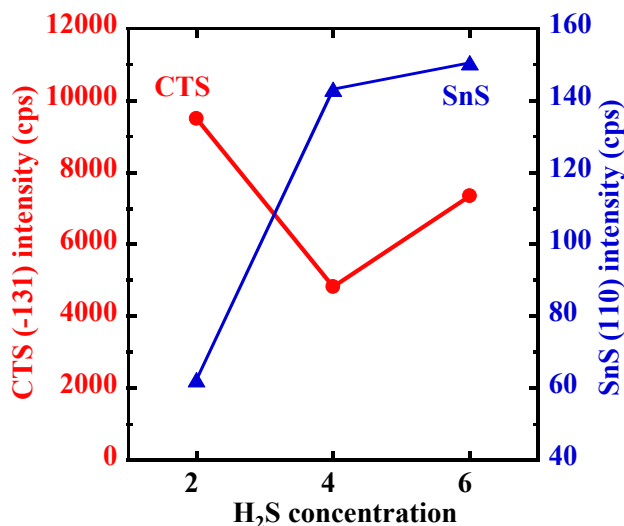


Fig. 2 CTS (-133) and SnS (110) intensity of CTS thin films prepared at different H₂S concentrations

Fig. 3 に H₂S 濃度を変化させた時の組成比を示す。H₂S 濃度の増加によって、Sn および S 組成は減少し、Cu 組成は増加するという結果になった。この結果より、アニール中の H₂S 濃度の増加は、膜中の S 組成の増加に繋がらないということが分かった。また、Sn と S 組成が減少していることから、H₂S 濃度の増加により、SnS 蒸発が増加していると考えられる。さらに、プリカーサ時の Cu/Sn 比は 1.75 であるのに対し、アニール後の全ての薄膜で Cu/Sn 比は 2 以上となった。

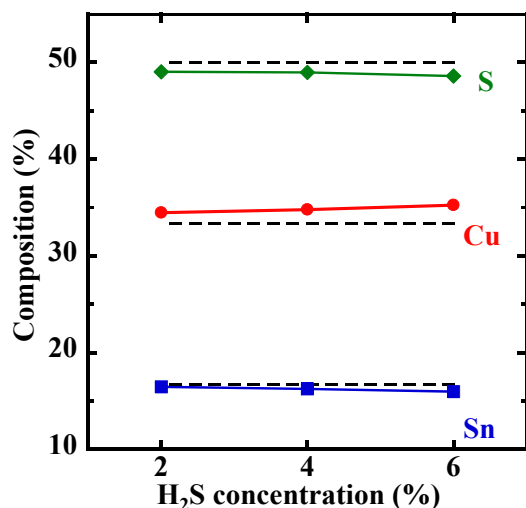
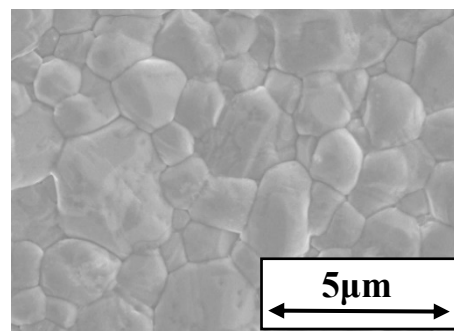


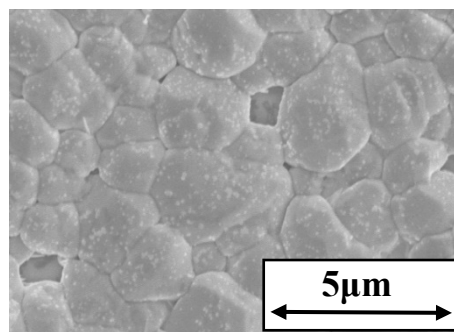
Fig. 3 Composition of CTS thin films prepared at different H₂S concentrations

Fig. 4 に異なる H₂S 濃度で作製した CTS 薄膜の表面形態を示す。全ての H₂S 濃度において 1~2.5μm 程度の粒径が得られ、変換効率 5.1% を達成している薄膜の結晶粒径よりも 2 倍以上の結晶を得ることが

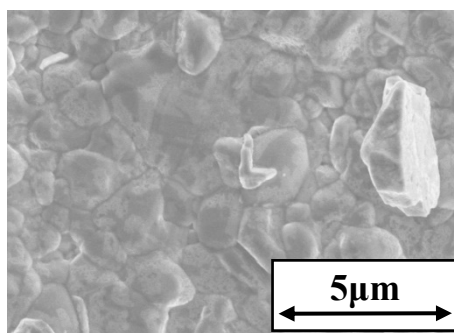
出来た⁴⁾。この結果より、H₂S アニールが固体硫黄を使ったアニールに比べて結晶成長に効果的であることが分かった。特に 2% では、密で平坦な膜が形成されており、小さい結晶粒と表面の荒さを改善することが出来た。しかしながら、4% では、SnS の再蒸発によって形成されたものだと考えられる多数のボイドが確認され、6% にすると、ボイドに加えて異相に起因したと考えられるものも確認された。



(a) 2%



(b) 4%



(c) 6%

Fig. 4 Surface SEM images of CTS thin films prepared at different H₂S concentrations

Fig. 5 に H₂S 濃度を変化させた時の電気的特性を示す。H₂S 濃度を増加させると、キャリア密度は増加し、移動度は減少する傾向が見られた。2、4% において、キャリア密度は $2\sim 3 \times 10^{18} \text{cm}^{-3}$ であり、変換効率 2.9% を達成している薄膜と同等の結果となった⁶⁾。また、6% にすると、キャリア密度は 1 桁増加し、比較的高いキャリア密度になった。これは、XRD

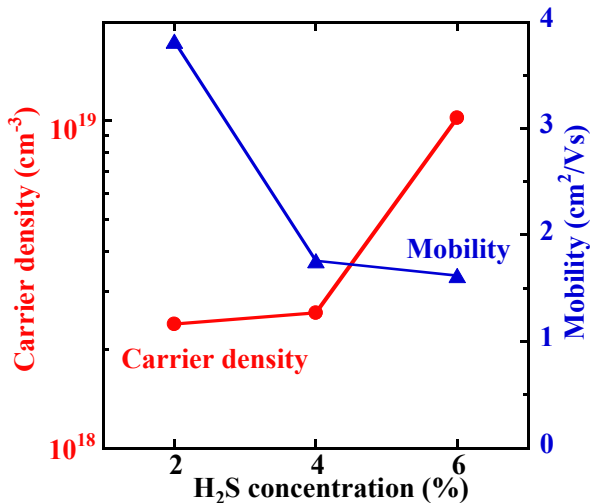


Fig. 5 Carrier density and mobility of CTS thin films prepared at different H₂S concentrations

や表面 SEM 像で観察された異相により増加したと考える。さらに、2%の時の移動度は、変換効率 2.9% を達成している薄膜に比べて 5 倍以上大きい結果となった⁶⁾。これは、大粒径でボイドのない薄膜を得られたことが要因だと考える。結果的に、2%において適切なキャリア密度と高い移動度が実現でき、高品質な膜が得られたと考える。

3.2 CdS 薄膜における膜厚の検討

Fig. 6 に CdS 薄膜の成膜時間を変化させて作製した太陽電池デバイスの断面元素マッピングを、Fig. 7

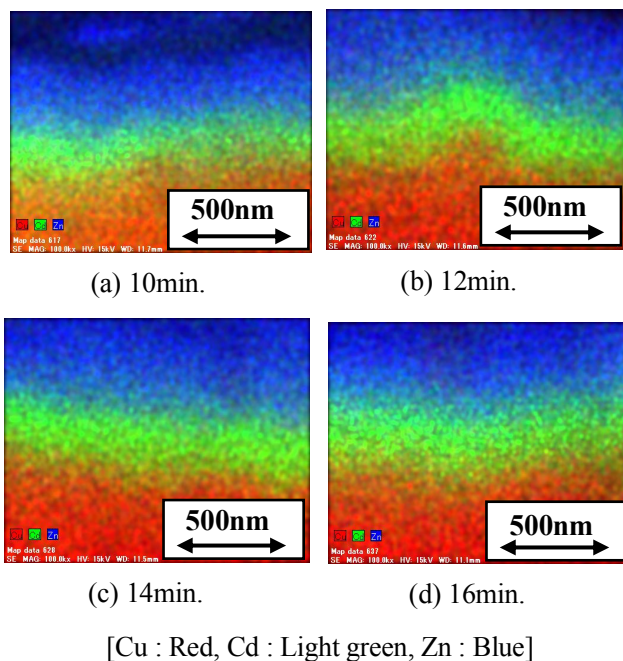


Fig. 6 Cross sectional mapping of CTS thin film solar cells with varying deposition time of CdS thin films

に断面元素マッピングから推定した CdS 薄膜の膜厚を示す。断面元素マッピングより、CdS 薄膜は荒い CTS 薄膜の表面を覆うように形成されていることが確認できた。また、成膜時間を長くすると、膜厚は顕著に増加しており、成膜時間による膜厚制御が可能であることが分かった。

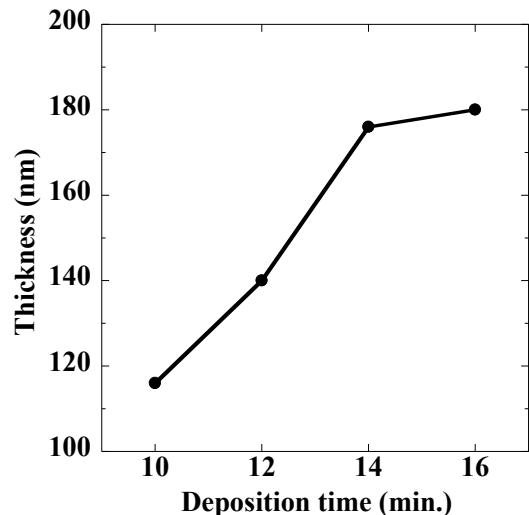


Fig. 7 Cross sectional mapping of CTS thin film solar cells with varying deposition times of CdS thin films

Fig. 8 に CdS 薄膜の成膜時間を変化させて作製した太陽電池デバイスの I-V 特性を、Fig. 9 に太陽電池特性を示す。成膜時間を長くすると、変換効率は向上する傾向が見られ、14 分の時に最も高い変換効率 1.16% が得ることができた。しかし、16 分にすると変換効率は減少した。この時、10~14 分の成膜時間によって最も変化が見られたパラメータは開放電圧と曲線因子であり、短絡電流密度は大きな変化

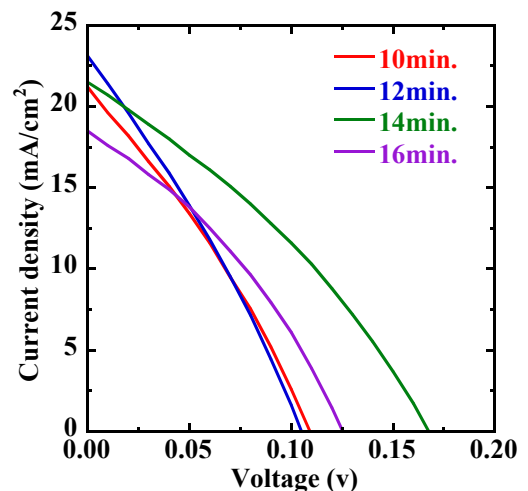


Fig. 8 I-V curve of CTS thin film solar cells with varying deposition times of CdS thin films

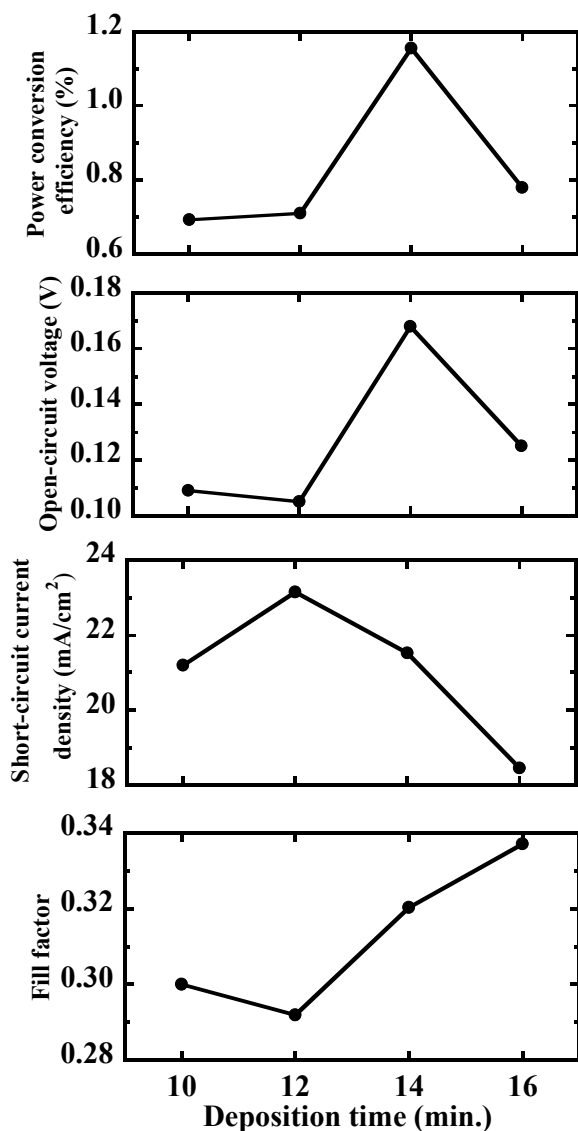


Fig. 9 Device performance of CTS thin film solar cells with varying deposition times of CdS thin films

は見られなかった。これは、膜厚が厚いことで、漏れ電流が流れるのを防いだことが要因であると考えられる。漏れ電流に影響する並列抵抗を計算すると、10分では 37.9Ω であり、14分では 64.3Ω と2倍近く増加していた。並列抵抗が大きいほど漏れ電流は小さくなるため、膜厚増加によって漏れ電流が減少していることが裏付けられた。しかしながら、16分では並列抵抗が14分の時と変わらないが、変換効率が減少する結果になった。これは、成膜時間が長いことでCTS内にCdが拡散し、再結合中心になったことで少数キャリア拡散長が減少したと考えている。

Fig. 10 に CdS 薄膜の成膜時間を変化させて作製した太陽電池デバイスの外部量子効率を示す。10分から14分と成膜時間を増加させると、500~1300nmの広い波長範囲で外部量子効率が徐々に改善した。

しかしながら、16分にすると低下した。この結果は、太陽電池特性の結果と一致しており、成膜時間が外部量子効率に影響を及ぼすことが分かった。

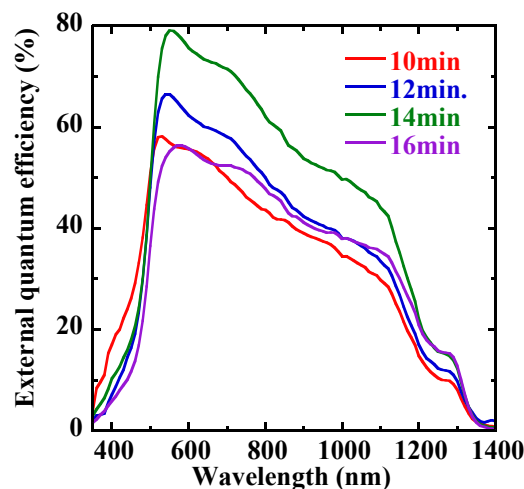


Fig. 10 EQE of CTS thin film solar cells with varying deposition times of CdS thin films

4 まとめ

CTS 薄膜の更なる高品質化を狙うために、H₂S 濃度が薄膜に与える影響について調査した。XRD より、H₂S 濃度 2%において、SnS ピーク強度が最も減少し、単相に近い薄膜が得られた。組成分析より、H₂S 濃度増加に伴い、S および Sn 組成が減少し、Cu 組成が増加したことから、SnS 蒸発が顕著に増加している可能性があることが分かった。表面形態より、全ての薄膜において1~2.5 μ m程度の粒径が得られ、最高変換効率を達成している薄膜よりも2倍以上の粒径を得ることが出来た。ホール効果測定より、2%において、高い変換効率を達成している薄膜のキャリア密度と同等の値が得られる一方で、5倍以上の高い移動度が実現できた。これは、大粒径でポイドのない薄膜が得られたことが要因だと考える。これらの結果から、H₂S 濃度を2%において最も高品質な薄膜を得ることができた。

また CTS 太陽電池の更なる高効率化にむけて、CdS の成膜時間を変化させ、膜厚の検討を行った。その結果、成膜時間14分において最も高い変換効率1.16%が得られた。これは、膜厚増加によって漏れ電流が抑制された結果、曲線因子と開放電圧が改善され変換効率向上に繋がったと考える。

参考文献

1) Titilayo A. Kuku, Olaosebikan A. Fakolujo:

- Photovoltaic characteristics of thin films of Cu_2SnS_3 ,
Solar Energy Materials, 16, 199-204 (1987)
- 2) S. P. Bremner, M. Y. Levy, C. B. Honsberg : Analysis of tandem solar cell efficiencies under AM1.5G spectrum using a rapid flux calculation method, *PROGRESS IN PHOTOVOLTAICS*, 16, 225-233 (2008)
- 3) Jakapan Chantana, Kanta Tai, Haruki Hayashi, Takahito Nishimura, Yu Kawano, Takashi Minemoto : Investigation of carrier recombination of Na-doped Cu_2SnS_3 solar cell for its improved conversion efficiency of 5.1%”, *Solar Energy Materials & Solar Cells*, 206, 110261 (2020)
- 4) Vasudeva Reddy Minnam Reddy, Mohan Reddy Pallavolua, Phaneendra Reddy Guddetib, Sreedevi Gedia, Kishore Kumar Yarragudi Bathal Reddy, Babu Pejjaia, Woo Kyoung Kima, Thulasi Ramakrishna Reddy Kotteb, Chinho Parka : Review on Cu_2SnS_3 , Cu_3SnS_4 , and Cu_4SnS_4 thin films and their photovoltaic performance, *Journal of Industrial and Engineering Chemistry*, 76, 39-74 (2019)
- 5) 中田時夫、日刊工業新聞社、CIGS 太陽電池の基礎技術、p106
- 6) Naoya Aihara, Hideaki Araki, Akiko Takeuchi, Kazuo Jimbo, Hironori Katagiri : Fabrication of Cu_2SnS_3 thin films by sulfurization of evaporated Cu-Sn precursors for solar cells, *physic status solidi c*, 10, 1086-1092 (2013)

資料・紹介論文編

ワープロ演習での文集作成の一実施例

中村博文¹

An Example of Creating Collections of Compositions in Word Processor Exercise

NAKAMURA Hirofumi¹

(令和3年10月1日受理)

あらまし ワープロ（ワードプロセッサ）演習を伴う授業の一部でクラス毎の文集を作成した一実施例を報告する。クラスメンバー各自が書いた原稿を収集し、必要数を印刷して、簡素でも製本したものが各自に残るという点で他の文集と変わりはないが、科目の内容、原稿執筆の負担、準備などの兼ね合いの中でこれまで授業のワープロ課題のひとつとして実施してきたものである。授業ではミニ文集と呼んでいる。

授業の演習課題の扱いであり、指示や連絡事項や必要なワープロ操作の内容をプリントにして配布し、下書きは宿題として、更に、ある程度の時間以上の入力や設定のワープロ作業も宿題として実施している。その配布プリントには、執筆する上での留意事項や、筆が進まないケースへのテーマ例も載せている。先々見返すときに過去のことばかりにならない執筆内容をいくらかでも設けたいと考え、自由題での執筆の他に、そのような指定の執筆事項もいくつか設けている。原稿として、製本用の印刷原稿として使える A4判 1 ページで印刷したものを提出することとしている。その原稿で必要数の印刷をした後、要領を説明して、更に、簡素ながらも各自で自分のものを製本することとしている。けがや作業のミスがないことは勿論のことであるが、使いやすさという点にも触れている。他に、ミニ文集作成を実施してきた中での進行上の非原則的な場合への対応例についても触れる。

在学時のものではあるが学生の感想としては、例えば、楽しんで執筆した、当該科目の一番きつい課題であったなどがある。文章を書くことの得手不得手との関連性は大きいようである。

キーワード [文集, 授業, 演習, ワードプロセッサ, 製本]

1 はじめに

本稿は、ワープロ（ワードプロセッサ）演習を伴う授業の演習課題のひとつとして、クラス毎に文集作成をしてきた一実施例の報告である。授業ではミニ文集と呼んでいる。

本稿でクラスメンバーのことを本校での通常の言い方である「学生」と表記するが、本稿の内容について生徒との違いはない。既学習や経験の差以外では児童でも同様である。大部分で「クラス」と表記

するが、本校のクラスは学年ごとには組ではなく定員 40 名の学科で分かれている。2 学期制である。

執筆したものを収集し、必要数を印刷（以下、「製本用印刷」と表記する）し、簡素でも製本したものが執筆者各自に残るという点は、多くの文集と違いはない。決して継続的にスマートに芳しく実施できている取り組みなどとは言えないが、科目の内容、学生の負担、準備などの兼ね合いの中でこれまでワープロ課題のひとつとして実施してきたものである。変更を重ねてきているが、それに多くは触れず、最

近の状況を主として述べる。試行錯誤も実施できなかった年度もありながらの四半世紀余りの変遷を多くは記さないため、卒業生、在校生によっては異なる指示や進行を経験している場合もある。その場合に、無視されたなどとは思わないで“進化”したと笑覧して頂けたら幸いである。

多くは言い切る言い回しで述べるが、状況によって、公平性は意識しながらも本稿に記載しきれない個別の対応をしている場合もある。

筆者の状況特有の事柄だけでなく、どなたがどこで実施されても大差ない内容でも、実施のために必要なことであれば例として記している。

実施の具体的な事柄に関して同僚や学生からの意見を取り入れている事柄があるが、正確に記憶していないため提案者等については記載を省く。また、有益な意見を聞きながら取り入れられていない場合もある。容赦頂きたい。

進行において、病欠その他による学生の遅滞などはいつでもあり得る。対応を、まとめて一節にもできるが、それでは原則的な進行の個々の記述から遠くなるため、関連する所で述べている。

2 ミニ文集について

ミニ文集作成を実施するようになったきっかけについて2.1で、ミニ文集と呼んでいることについて2.2で述べる。更に後ろに、多くの文集と多少異なるのではと筆者が思っている事柄も含め、以下の内容、順番で述べる。

- ・学生に実施内容の概要を伝え、下書きを宿題という形で準備に入ってもらおう（初期に伝達は口頭のみであったが、その内に関連事項やワープロ操作なども含めて、プリント配布するようになった。その内に「実施要領」と呼ぶようになった。実施要領に関しては付録1で述べる）。
- ・学生は、自由題またはひとつの指定題ではなく、2.3に述べるように、自由題と複数の指定題で執筆することとしている。
- ・各自の持ち分A4判1ページ分の原稿（以下、単に「原稿」と表記したときは学生が作成する原稿のことである）の作成にワープロを用いることとしている。ワープロ演習のひとつとして2.4で述べる演習内容を含んだ設定等を行うこととしている。そして、製本用印刷でそのまま使える印刷したものを原稿として提出することとしている（以下で「書く」や「執筆」と表記するが、手書きの下書きだけでなく、ワープロ入力を含む）。

- ・提出遅れの場合や寄稿も含めた、冊子としてのミニ文集の構成を2.5に述べる。
- ・製本に先立って原稿をもとに必要な枚数を印刷する作業（製本用印刷のこと）は2.6で述べる。なお、使用する印刷機は学生が使用するための機器ではないため、製本用印刷に学生は関わらない。
- ・製本は、簡素な製本（以下、簡易製本と表記する）ではあるが、各学生が自分の分は自分で製本することとしている。2.7で述べる。
- ・ミニ文集はクラス毎に作成するもののクラスの企画ではない。担任や、次年度の1学年上の同学科クラス担任（3月頃の配布時点で正確には予定者であるが、以下では単に「次年度担任」と表記する）への、配布や連絡について2.8で述べる。
- ・以上の実施の時期や進行上のタイミングについて2.9で述べる。

更に、特に異なった進行になった年度について3に、聞いた感想等を4に述べる。カレンダーの付属について付録2に、懸案の事項等を付録3に述べる。

2.1 きっかけ

本校で文集の作成はどの学年の活動内容にも入っていない（知る範囲で担任主動で作成された例が複数ある。学寮は卒業の文集を毎年度作成している）。

本校で情報リテラシー科目が始まる（全学科共通科目は平成10年度。ミニ文集を4学科とも作成したのはこれ以前）より前に、コンピュータプログラムによる問題解決を中心とした科目で、ワープロ演習も含めていた時期（記録はないが平成7年度以前）があった。ワープロ入力の内容を一部自由題にしたことがあり、更に、学生によってはそれを楽しむケースも見られたが、他の課題同様、印刷して提出すれば終わりであった（同じものを学生自身が永く保管したケースがあるかどうかは不明である）。

そういう中で、更に、製本用の必要枚数の印刷と、簡素でもホチキス（ステープラー）と製本テープによる製本を足せば、演習物が学生の手に残ることに思い至った。全体としては、操作の一部は学習して欲しい事柄であり授業時間もいづれか充てられそうであったこと、学生が宿題として下書きやワープロ入力の一部ができそうであったこと（自宅等でワープロが使えなくても他の宿題と同様にパソコン演習室、以下では「演習室」、で可能）、進行に応じた筆者の必要な対応（原稿の内容確認、製本用印刷、簡易製本の準備など）ができそうであったこと、クラス毎に約千枚の印刷や簡易製本に関する消耗品等の費用が大きくはないと思われたことなどである。

2.2 呼び方について

最初は「文集」と呼んでいたが、途中から「ミニ文集」と呼ぶように変えた。クラス等で文集を作る計画が出てきた場合(実際にそういうことがあった)にそれと言いかち合うことを最初から避けたかったことや、学生がより気楽に取り組むことを期待したことや、製本が簡素であることからである。

しかし、決してミニではない或いはミニではなかったと話す学生もいる。省かれていることがあるわけではないためそれは無理からぬ感想である。

ミニ文集と言っても、文集という言葉は入っているため、通常の課題とは異なり後に残るものを作る意味合いは抜けていない。

2.3 ミニ文集への執筆内容

自由題の他に、実施の年数を重なる内に、指定題もいくつか設けるようになった。近年は、自由題(700字以上)の他に、次の事柄も書くこととしている。数十年後でも内容が古くならずに見られるような事柄も、近年は少し含めている。

それは、自由題の後に順に8項目で、自己紹介、高専を選んだ訳と今の学科を選んだ訳、これまでまたは最近のこと、将来就きたい仕事、これからの自分について、将来の自分へのメッセージ、将来の自分の子供か子の世代へのメッセージ、今後50年以内に起こる事柄・変化・発明の予言・予想である(将来就きたい仕事が200字以上、他は100字以上としている)。

内容はオリジナルなものに限るとしているが、他の発表機会の内容と重なって構わないとしている。

取りかかりやすいよう、下書き用に用紙を1枚配布している。下書きでは記述の順序は問わないこととしている。手で筆記しないで直接ワープロに入力しても構わないことも伝えている。

2.4 ワープロ演習としての側面

近年は以下のような内容を含んでいる。

- ・持ち分 A4 判 1 ページの中で 1 段組み(氏名等において)と 2 段以上の段組み(本文において)とを使うこと。
- ・ルビを振ること(氏名において)。
- ・字数の確認をすること(内容毎に最少字数を指定している)。
- ・校閲機能を使うこと。
- ・原稿(サイズは、印刷が手軽な A4 判)を演習室外(校内共用施設、コンビニエンスストア、自宅や寮など)で印刷すること(以前、演習室にプリンタが設置されていた頃は、演習室や自宅等で印

刷)。

他に、この課題に限らないが、余白や行間の設定、印刷プレビューの活用、必要な PDF ファイルの作成を含み、段落の最初での桁下げを含んでいる。

縦書きか横書きか混成かは学生に任せており、縦書きや混成の方法も提示している。フォントはいろいろと試せるようどこでも自由としている。文字サイズも 10 ポイント以上で自由としている。

縦書きやイラスト等を添えることは必須とはしていない。これらは若干の加点事項としているものの、取り入れる学生は多くはない。

学生は指定した日時までに原稿を提出すれば原稿作成は一応終わりである。但し、中には、指示した設定等が未了の場合や、将来も含め他者の目に触れるのがはばかれるような内容もあり、筆者が目を通していている。場合によっては再提出を求めている。

指示した要領に従って文書をまとめる演習でもある。禁止事項等や誤字脱字は各自の責任でチェックすることとしている。誤字脱字はあまり指摘しないが、他のことは状況によってやり直しを求めている。

2.5 ミニ文集の体裁や構成内容

2.5.1 冊子としてのサイズと体裁について

校内でプリント印刷に通常用いられている用紙のサイズは A4 判と B4 判である。白のやや薄い用紙と、それよりやや厚めのやや茶色の用紙がある。

ミニ文集では B4 判の白の用紙の片面に 2 人分の A4 判原稿を縮小印刷している(ごく初期に他の印刷の仕方でも実施したかどうかは不明である)。表紙なども印刷機用の原稿は A4 判で作成し、同様に印刷する。

簡易製本時に折る手間が生じるものの、1 回で 2 人分を印刷できることは利点である。出来上がりのサイズがコンパクトでもある。

横書きの原稿が多いが、縦書きが増す期待から、近年は右綴じである(初期から長い間左綴じ)。

2.5.2 構成内容

ミニ文集のページ構成は、順に、表紙、表紙の裏、学生の執筆内容(寄せられたときは更に寄稿も)、空きページができる場合にメモ用ページ、筆者が書き込んだ 1 ページの情報関連等記事、筆者が付けた場合に見開き計 2 ページのカレンダー、裏表紙である。

以下、これらのそれぞれについて述べる。

2.5.2.1 表紙

表紙は、筆者が作ったものを製本用印刷の原稿にしている場合が殆どである。その表紙に「ミニ文集」と記す時期が途中続いたが、近年は、ミニという字を付けなくてもよいのではと思い、表紙では「文集」

にしている。この他に年度、学科名、学年も入れている。いつからか、授業科目名と簡易製本した年月（和暦と西暦）も入れるようになった。

2.5.2.2 表紙の裏

表紙の裏は、従来空白であったが、近年、学生のことだけに限らず、見る人の、取り扱いの注意を促す少量の文を入れている。

2.5.2.3 学生の執筆内容と寄稿

寄せ書き（手書き可）なども入れられると伝えているが、これまでない。

学生の手稿は、原則、出席番号順である。

原稿提出の締め切りからいく分待って、最初の製本用印刷時点で間に合っておらず簡易製本の前日頃までに提出された原稿は、更に後ろ側に出席番号順で続く。寄稿があるとき、寄稿はこの後である。

ページ番号は振らないことにしている。

2.5.2.4 メモ用のページ

原稿 2 人分ずつを B4 判用紙に縮小印刷している（2.5.1）。原稿（あれば寄稿も）が奇数の場合に、できるだけ、片方を空白のページにはせず、「メモ」の 2 字だけ入れたメモ用ページにしている。偶数の場合に、特に基準はないが、メモ用ページは設けなかったり 2 ページ分（1 枚）設けたりしている。

2.5.2.5 情報関連等記事の記載

広く情報や科学などに関連したニュースのキーワード等を適当に選択して載せている（近年、未来の予定も載せるようになった）。

2.5.2.6 カレンダー

必須のものではないため毎年ではないが、140 年分のカレンダーを見開き 2 ページで、これだけはレーザープリンタで出力して付けている。この他のことは付録 2 に述べる。

2.5.2.7 裏表紙

裏表紙は何も印刷していない。

2.6 製本用印刷

この作業に手落ちがあるとせっかく簡易製本をしても内容に不都合が生じるため、必要な原稿を正しい順序でズレやシワがなく必要数印刷できたかの確認は重要である。

製本用印刷は、廉価であることから、通常の配布プリント用の印刷機によるモノクロ印刷である。学生へは普段の配布プリント程度のにじみやかすれが生じることや、印刷の際に字の太さが変わる可能性も知らせている（詳しくは伝えないが、現在使用している印刷機は写真と文字の両対応の選択をすると文字が細る。原稿中の写真によってはこの設定でなくうまく印刷されない場合がある）。

2.6.1 製本用印刷の枚数

学生、配布教員、筆者（学生と同様に簡易製本する確認用）の分が必要枚数である（筆者がその年か次の年の担任であるクラスは 1 冊分減らす）。印刷時の印刷マスター作成の際に自動的に 1 枚印刷されるものを、ある程度読めれば、筆者用に充てている。

簡易製本時のトラブルで必要になった場合に、1 枚分は筆者の分を充てたり交換して対応する。例えば、作業中の破損や、誰かが 2 枚取っているなどで足りないときに直ぐに分からない場合などである。

初期はトラブル対応への予備として更に 2 枚や 1 枚ずつ余分に印刷していたが、なくても足りるケースが多く、それはしなくなった。

まれに（左綴じの頃にも）綴じる場所を間違える学生がおり、見えやすい面である表紙と裏表紙の用紙は交換できるものがあれば交換している。表紙を含む用紙（裏表紙を含む用紙が印刷の場合はそれも）は 1 枚程度余分に印刷しておくこともある。

印刷時間はクラスあたり近年は 30 分弱である。

2.6.2 原稿提出の遅れがある場合

遅れている原稿は、口頭や出席番号をクラス掲示して提出を促し、最大限待つが、製本用印刷の大部分は念のため簡易製本の数日前には終えている。その時点で全員の原稿が揃えばひとクラス分の印刷作業を 1 回で終わられる。しかし、その時点でも未提出者がいることはよくあり、その場合、簡易製本の前日頃にそれまでの提出分を追加印刷している。

1 回目の印刷に間に合った最後の出席番号の学生の手稿（寄稿がある場合は寄稿）が奇数人目の場合は、残しておいて、追加印刷の最初（寄稿の場合は最後）に位置させて印刷している。他で一言で製本用印刷と表記しているが、必要なこの追加印刷を含む。

以前より学年末試験後に答案確認をする授業時間があり、遅れて試験直後までに提出された原稿は（製本用印刷とは言い難いがここで述べる）、簡易製本したミニ文集に挟めるように準備（製本用印刷と同様に印刷して、その後で左右を十数ミリカット。山折りにしてミニ文集にはみ出さずに挟める大きさ。1 人分のときは原稿を複製して半分の印刷枚数で同様のものを切り出すこともある）してクラスに配布している。担任へも準備する。

2.6.3 印刷の不具合

印刷したものにズレやシワがあることがある。確認はするが、簡易製本時に初めて分かる場合もある。1 人分なら筆者のものと交換している。複数人分の場合は、内容が分かるなら勘弁してもらっている。中には、簡易製本時に若干名分を複写機で複製（原

稿を縮小コピー)して対応したこともある。

2.7 簡易製本

学生の安全と出来上がりの使い勝手に関して、最も留意すべき作業である。

2.7.1 簡易製本の手順

学生が自他共にけがのないことと使い難いものを作らないようにという点を含め、下記の手順等を伝えて(若干の図示を交えている)作業に入っている。なお、抜針器を用意しているので綴じ方に不都合があったときは直ぐ知らせること、ハサミや平口で自他ともにけがをしないようにということも伝えている。ホチキスの針を素手で外そうとしてけがをした学生がいたことを話すこともある。

- ・指示した順に続いて席を立って、指示した動線で移動して、予め筆者が印刷した 20 数種類の用紙(表紙を含むものから裏表紙を含むものまで印刷した、40 数枚毎の束)を並べて(各束を縦長で横並びで)置いた所から印刷物を順に 1 枚ずつ全て拾い集めて(白紙や不都合なものは近くによけておいてもらう)、指示した動線で自席に戻る。
- ・用紙をすべて山折りにする(角同士のズレは 0.5 mm 以下を目指す。慣れると 3 枚まではいっぺんに折ってもある程度の精度で折れる。山の部分が丸まらないように折る。仕上げに物でこすってもよいが、机をこすらないこと。爪でもできる)。
- ・順序、向き、枚数を確認する(枚数を伝えておくのでそれを確認する。印刷の不都合や過不足がある場合は筆者に知らせる)。
- ・器具の所に移動する(2 か所の器具等の場所と、各自に色を選んでもらうため置いている製本テープの色と起き場所も伝える)。
- ・用紙の束を揃える(4 辺とも。完成時に指をあててパラパラめくりができるのが具合がよく、そのためには、製本テープで隠れる背ではなく、その反対側を最後にきれいに揃える)。
- ・ホチキスで 3 か所綴じる(中央と、端から 2 cm 位の所。ホチキス自身の台が狭いためふくらみができないよう留意。右綴じ以外の綴じ方をしないよう留意。誰かが押してあげるとよいものができやすく全体が速くなる。ホチキスの台のストッパーが使う毎にずれていくことに留意。せかさない。綴じ方の異常や針が切れた際は筆者に知らせる)。
- ・閉じた後のホチキスの針の盛り上がりや、平口で挟んでつぶして埋没させる(盛り上がったままだと先々貼った製本テープを針が突き破るため。隠れるので、その部分の見かけは悪くてもよい。斜めに倒してつぶさないよう留意。使い終わった平口は、机に置かず、箱に戻すか手渡しで渡す)。
- ・使いたい色の製本テープを選んで、ホチキスで綴じた冊子にあてがって、必要な長さをハサミで切り出す(後で余りを切り揃えるのではなく、最初に長さ分切り出す。使い終わったハサミと製本テープは、机に置かず、箱に戻すか手渡しで渡す)。
- ・自席に戻る。
- ・製本テープを貼る(例えば、台紙の半分をはがして貼って、断面がコの字型でなくユの字型になるように折り曲げながら癖をつけて、残りの台紙をはがして、端から端まで少しずつ貼るつもりで。表側と裏側の製本テープを普通同時には見ないため、厚みを考慮せず表裏で幅が異なっても気にしなくてよい。端を片方貼ってから、他方へひとなべて済まそうとすると、傾きやすい。一度貼るときれいにははがせない。触れた程度なら可)。
- ・製本テープの台紙を指定したゴミ箱に廃棄する。

作業の完了が早い学生も多いが、説明を始めて最後の学生の終了までに 4, 50 分かかることもある。

作業中に読むと時間が伸びる。順序等の確認以外は出来上がるまで内容を読まないよう指示している。

印刷したものを学生が取り終えた後で回収したのは配布教員と筆者の分である。現担任の分の簡易製本を学生に頼むこともあったが、多くは後で不器用ながらも筆者がしている。

2.7.2 器具等の準備

準備するのは、ホチキス(針が長さ約 1 cm のもの。針は減っていたら補充)、平口、ハサミ、道具や製本テープを入れる箱、ゴミ箱(ここまで 2 セット)、製本テープ(35 mm 幅。黒は道具等 2 セットの両方に配置、他は色の系統で適当に 2 分して配置)、補充用のホチキスの針と製本テープ、抜針器である。

製本テープの色は、初期は黒のみで、その内に適当に 5 種類程度を加えるようになった。減りが一番多いのは黒であるが、他の色も使われている。

2.7.3 印刷した用紙を拾い集める動線について

簡易製本の全作業を演習室内で済ませている。印刷した用紙の拾い集めは、人流の衝突が少ない動線を指定する(難しい場合は、20 数種類を通常の配布物と同様に配布することもある)。概ね演習室のパソコンの列毎順であるが、中には、学生が着席している机に印刷した用紙を置かざるを得ない場合があり、その場合、クラス全体のための協力ということで、その席の学生 10 人弱に動線に関係ない場所で立って待ってもらって、最後尾として巡ってもらう。その学生達に続いて筆者が残りの用紙を回収する毎に、

その学生達が順に自席に着席する。

2.7.4 欠席者への対応について

欠席学生の分については、可能な場合は隣席の2名の学生に頼んで、分担して簡易製本して（特に20数枚の山折りは時間を要するため2人で分担）、渡してもらっている。また、お願いした簡易製本が終わるまでは授業は進めずに待つことを伝えている。

製本用印刷した用紙の拾い集めは、2人のうちひとりの学生が自分の分と欠席学生の分の2回巡る必要があるため、少しでも時間短縮になるよう、クラス内に、順序を待つことなく最優先で列に入れる特権を認めることを周知している

欠席学生の分をその場で頼まない場合は、後で筆者が簡易製本して渡している。欠席が続いているなどの学生の対応は担任に相談することもある。

2.7.5 学生の誤作業への対応について

ホチキスの綴じ位置を間違えたと分かった場合は筆者がホチキスの針を抜いている。用紙の交換については2.6.1で述べた。

ページの重複や抜けも製本テープを貼る前に分かれば筆者がホチキスの針を抜いて対応している。

ごくまれに他の学生のミスに反応する学生がいる。中にはミスをした本人が反応を誘っている場合もある。どちらも悪気のないものもある。傷つけるような言動が広がらないようには気を付けている。

2.7.6 ホチキスの針の長さ

透けない用紙に計4ページ分を両面印刷して折った側を綴じる方法も考えられるが、多くの学生で折り曲げた後にページ毎のズレが小さくない（指をあててパラパラめくりができないほどの凸凹がある）仕上がりになることが心配されたため、仕上がりにより安定する、計2ページ分を片面印刷して山折りにして山と反対側で綴じる方法を選んだ（山の側は揃えやすいため）。

こうすると、薄い印刷用紙でも長さが5mm位の通常のホチキスの針では50枚程度分の厚みを十分綴じ切れないため、ホチキスは針の長さが約1cmのものを用いている。

2.8 配布範囲や担任への連絡

実施を重ねている内に、クラス担任にも差し上げるようになった。その後、作成完了時期がほぼ年度末であることや、自己紹介なども記載してもらっていることから、次年度担任にも差し上げるようになった。学生にはこれらについて取りかかり前に可能性を伝えている。

クラスで文集などを企画される可能性はあるため、クラス担任にミニ文集を作るあらしを秋頃に連絡

している。寄稿の検討依頼や、寄稿を頂く際の書式等も連絡している（上限を1ページとしたことも設けなかったこともある。寄稿がない年度の方が多いが、4クラスの内1クラスほど頂くことがある）。

担任へは、寄稿を頂いた場合に筆者から学生に簡易製本前には話さないことを伝えている。また、次年度担任への配布分に寄稿が載るのを遠慮される場合は連絡を頂くことにしている。

クラス担任にミニ文集について連絡をしていて、クラス担任はご存じのため、学生原稿提出の促進などで配慮を頂いている場合もある。

次年度担任には念のため、ミニ文集のことと、見ない方が学生と接しやすいなど無用の場合のみ返信をと、製本用印刷の前頃に連絡している。受け取られなかったことはない。その連絡の際に、ミニ文集の所持状況や会話内容で次年度の担任を学生に始業式前に悟られないように言葉を添えている。

2.9 実施の時期や進行のタイミング

2.9.1 実施学年や時期

科目の内容や学習時期との関連から、1年生のワープロ演習の後期半ば以降がミニ文集の入力時期になるように実施している（初期は2年生で異なる時期の場合もあった。「数理・データサイエンス・AI教育プログラム認定制度（リテラシーレベル）」への本校の認定申請に伴い授業内容の改訂が検討中であり令和4年度以降のミニ文集の実施は未定である）。

2.9.2 進行のタイミング

ミニ文集の実施の伝達は、10月頃に実施要領のプリント配布を伴ってしている（近年は後期中間試験が12月初め頃であるが、比較的初期に後期中間試験が12月半ば以降の頃は12月に伝えて下書きを冬休みの宿題にしていたときもあった。本校の秋の体育や文化の行事終了後の11月初めにしていた時期もあったが、伝達は一層早めるようになった）。

ワープロ入力を始めるのは、後期中間試験明けの最初の授業時（近年は12月上旬頃、比較的初期は冬休み明け）に多少の時間を設けるときからで、その時までの課題としておいた下書きの完了や進捗の確認もする。その次の週も入力時間を設け、遅れている学生の確認もする。下書きが遅れていると、最初からワープロに向かって取り組もうとする学生もいる。参考まで全員に、中には、ワープロだと修正が多く、かえって書いて添削してからの方が、全体として仕上がり早い人もいることを伝えている。

ワープロ入力の時間をとった際に、印刷の仕方を伝えておく（この時点で入力や印刷まで自宅等で終えている学生もいる。初期は演習室での印刷時にパ

ソコンのフリーズがよくありデータ保存を促した)。

印刷した原稿の一応の提出締め切りは、学年末試験(近年は2月上旬頃、以前は2月末頃)の3週間前に設定している。病欠等で多少遅れるとか、再提出になるとかの程度なら、学生の作業が学年末試験の1週間頃には終わる目安で進めている。数学などの1月実施の実力試験の日程によって原稿の締め切りを加減することもある。取りかかった原稿を全員に提出してもらいたいが、学生に不都合の生じる場合もある。いつからか、最後は字数等が足りなくても受け取ると連絡するようになった。無理強いはしないつもりでいるが、提出の促し方の加減は悩ましい。

学年末試験より前の授業中に簡易製本を終えて各自がミニ文集を手にする。

3 異なる進行になった年度について

以上に述べた実施状況と特に異なった進行になった年度について述べる。

3.1 混合学級(平成9~13年度)

本校のクラスは学科毎であるが、本校は平成9~13年度の1年生に、専門科目は学科毎で、一般科目は別途年度初めに割り振った組毎で授業を受けるといって、混合学級を試行した。2種類のクラス編成のどちらも共に勉強したクラスであるため、ミニ文集は学科毎と組毎の両方を作成することにした(ミニ文集作成を実施していた科目は組毎のクラスで授業があった。記録していないが、初めは2種類ではなく組毎のみであったかも知れない)。

各自が作成する原稿の内容の多くは両方に共通でよいものの、例えば組毎のミニ文集に他学科の学生に無関係のことが載らないなどのためには、原稿も2種類執筆してもらうことが考えられた。しかし、筆者や学生の省力や物的な節約のため、原稿は1種類とした。それを2種類のミニ文集で使えるよう、氏名等の1行または1列の中に組名とその出席番号及び学科名とその出席番号を記載することとした。

製本用印刷において、提出原稿は、まず組毎の印刷で使い、整理し直して学科毎の印刷に用いた。

2冊分の簡易製本の作業は、混乱しないように、組毎と学科毎で時間帯をずらして行った。4組あるため、組毎の20数種類の印刷の束は組毎にほぼ全て簡易製本に回るが、学科毎の20数種類の印刷の束はそれぞれが約4分の1ずつ減っていく要領で進めた。

3.2 新型コロナウイルス感染症(令和2年度より)

新型コロナウイルス感染症が国内では令和2年1

月頃から報道され始めた。ミニ文集作成は、平成31年度は影響なかったが、令和2年度分は本稿の投稿時点でまだ製本用印刷と簡易製本が完了していない。

新型コロナウイルス感染症の影響で令和2年度に本校は遠隔授業で始まり、前期に演習室でのワープロ演習はできなかった。家庭でパソコンが使える学生もある程度いたことや、スマートフォンは皆が使用していたため、最終的に文字データとして揃えば後は筆者が補助して印刷原稿にできる見通しから、ミニ文集の作成をする方向で開始した(原稿ができれば最悪の場合に年度をまたいででも製本まで漕ぎつけると考えていたが、実際、年度を越している)。後期は12月まで全学年が原則対面授業であったため原稿入力を始めるタイミングまではほぼ例年通りに進んだ。この年に限らないが遅れている学生もいた。

1月から本校は遠隔授業になったため、印刷媒体による原稿提出を変更して、可能な学生にはPDF形式のファイルで提出することとした。他に、ワープロ形式のファイルや、遠隔授業システムのチャット機能を活用して文字データでも受け取ることにして、筆者がPDF形式の印刷原稿に仕上げた。どの場合も、学生が確認や修正や設定をすべき事柄は指摘して再提出を求めた。1回で済まない学生が例年より多かった。例年より他の演習が少なかったことや、学生の実施要領の確認不十分も含めて、遠隔授業システムで春休みまで個別に念を入れられた面もある。

簡易製本が新型コロナウイルス感染症の新たな感染源になってしまうことは避けたいため、この報告の提出時点でも、成り行きを見守っているところである。製本作業を筆者だけで行うことは不可能ではないものの、今のところは学生に自分のミニ文集は自分で製本する機会を設けたいと考えている。その際は、予め手順をまとめてデータまたはプリントで配布して読んでおいてもらって進める予定である。

令和3年度はミニ文集作成を実施しないこととした。後期のパソコン利用や、後期や来年度の簡易製本の時期の見通しが立たないためである。未製本の原稿が更に蓄積する懸念もあった。

4 聞いた感想等

思い出す範囲でミニ文集について学生(その時点で在学中)から聞いた感想等について記しておく。

楽しかった、課題ではあったけれども楽しんで書いた、級友の知らない側面を知ることができるという感想がある一方、科目の中で一番きつかった課題であったという感想もある。それぞれ、文章を書くのが好きだからとか苦手だからと理由まで聞いたケ

ースもある。存在自体を肯定や否定する意見もある。授業や宿題だから書くし製本もするけれども自分は読まないという学生もいた。

教員からの感想についても思い出す範囲で記す。

文章でまとめる機会のひとつになる、将来のことを考えるまたは確認する機会になる。

5 おわりに

授業のワープロ課題のひとつとしてクラス毎の文集を作成した一実施例を報告した。

それぞれの文集が後々ささやかでも良い意味で文集としての役をなすひとときがあるなら幸いである。収録に至らなかったページは追憶に任せたい。

付録1 実施要領について

初期には、ミニ文集に取りかかる前に数分の説明をして、後は個別に質問や提出原稿に対応していたが、大勢が同時期に取り組む際に同様の事柄をいく度もやり取りすることがあったり、説明を聞いていなかったり忘れていたり勝手に判断したりしていることもあったり、学習して欲しい内容は教科書との重複も含め記述した方がよいことなどから、プリント配布するようになった。授業では実施要領と呼んでいる。執筆要領に留まらない内容を含んでいる。

用紙への印刷要領（演習室にプリンタが設置されていた頃は、その関連した事柄を実施要領に載せていた）と簡易製本の要領は、別途連絡するため、実施要領には含めていない（近年はコンビニエンスストアでの印刷方法を2例ほどWeb上で提示している）。

実施要領は、いつからか、実施要領に沿った体裁で作成している。添削を重ねながらも少しずつ内容が増えて、現在は2ページ分になっている。

付図1に実施要領を例示している（余白の寸法を、本研究報告誌の投稿規定に合わせるために、88%に縮小して載せている。また、実際は、学生の便宜を図るために1枚目下部の少し外側に実寸の定規の画像が、2枚目下部に簡単なイラストが入っているが、同投稿規定に従ってこれらを含めると内容が更小さくなるため、付図1ではこれらを除いている）。

なお、付図1には次の実施を想定した修正を含んでいる。令和2年度と比べて、例えば、それまで太字の他に二重のアンダーラインも多用していたが圧迫感のあるケースがあり二重のアンダーラインは減

らしている。また、近年特に筆者のいろいろのプリントの密度が高いという学生の意見に対し（択一で尋ねた例を脚注^{*1}に記す）字数を減らし行間を2ポイントから2.5ポイントに25%広げている。今後の実施において3ページ構成にするかどうかは今のところ未定である。

実施要領の分量が学生の執筆量より多いが、削ると、その部分の質問があったり、質問がないまま勝手な解釈で原稿が提出されたりで、復活させたこともある。毎年添削はしているものの容易に削減ができないでいる。ある程度以上で仕上げたいなどの気持ちから削れないでいるものもある。

なお、実施要領の項目立てや内容や掲載位置は度々変わっている。

製本を右綴じでするようになってからも、実施要領は、半角の使いやすさから横書きのままである。

実施要領の内容を細かすぎると思う学生もおり、ごく最近では参考までそれぞれの学科の専門分野の関連学会の投稿規定の一部を提示したこともある。

以下、実施要領について簡単に説明する。多少の理由があるものにも触れる。

付録1.1 1のあらましについて

ここだけでなく、所々にフォント名を記している。一部の忘れて欲しくない事柄も最初に述べている。原稿提出だけで科目の評価が安泰というような誤解がないよう、近年、演習のひとつと明記している。

付録1.2 2の主な流れや日程について

充てる授業時間や締め切りの他に、各自の時間調整のために、他の日程との関連も述べている。

付録1.3 3の執筆内容や字数について

自由題（題を付けるかどうかは任意）の他にも執筆内容を指定している。将来において過去を振り返るだけではない内容や次世代に関連する内容も、近年いくつか設けている。

何を書いたらよいか分からないという学生がいたことへの対応として、途中から話題の例を挙げるようになった。

学生の執筆量を増やして合計で2千字程度で課した時もあったが減らしてきている。

ここに限らないが、書いて欲しくないことやして欲しくないことについて注意を与えている。ほとんどは、途中で改めさせた言わば未遂の場合も含め、

*1 令和2年度末に実施要領について41人と42人の計2クラスで、「丁度2枚分で配布していたミニ文集の実施要領①②に関してです。」に、聞き方に凸凹はあるが4択1問で尋ねて、39人と38人の回答の中で、「枚数が増えても、行間を空けるべきと思う」が15人と11人、「これだけはこれでしょうがないと思う」が14人と16人、「密度や枚数は今のでむしろ適切と思う」が8人と5人、「その他」が5人と5人であった。

過去にあった事柄がもとになっている。

住所、生年月日、電話番号などは書かないよう明記している。

実施を重ねる内に加点事項も設けるようになった。まれに英語での執筆もある。縦書きと英語での執筆は、両立が困難ということもあるが、それぞれ半分程度以上なら評価する。

付録 1.4 4 の設定・操作について

学習内容として実施して欲しい設定・操作や使う可能性がある設定・操作についてカバーするために、特にここには多く記している。中には沢山書きたい学生もいるため、文量が多くなった場合の、ミニ文集として許容するギリギリの設定も記している。

付録 1.4.1 余白寸法について

製本用印刷での印刷機にとっての原稿サイズは A3 版である (2.5.1)。これを A4 判の原稿 2 枚で構成するが、学生の原稿 (寄稿も) は、提出時点で左右どちらになるか予め確定できない。そのため、余白寸法は 1 種類にする必要がある。

筆者には左右が 2 対 1 の場合が扱いやすく、切りの良い寸法の、左は 30 mm、右は 15 mm としている。印刷時に片側に寄せて中央で原稿 2 枚を 15 mm 重ねるつもりで扱っている (筆者は印刷機に目印を置く場合も目分量の場合もある)。

余白寸法の許容を ±4 mm としている。±4 mm でなくても差し支えないが、簡易製本後に開いたときの見やすさと体裁の統一性からこの程度にしている。

付録 1.5 5 の縦書きへの補足について

縦書きにしない学生にも知っておいて欲しいため、縦書きにする場合の操作を記している。

付録 1.7 の内容もそうであるが、以前は他の内容と混ぜていた。独立して表記すると分かりやすい面はあるが、知っておいて欲しいと思っても、用いないとき読み飛ばすのをしやすくしている面もある。

付録 1.6 6 の氏名番号等について

学生にとっての確認になるよう氏名のローマ字や正確な学科名の表記も含めている。半角と全角の使い分けやルビ付けも練習する。

各原稿にはタイトルを付けないため、氏名番号等は最初の行または列に位置する (自由題中心の頃はタイトルを付けても構わないとしていた)。筆者にとって扱りがよいため、名前等の行は 1 段で出席番号を最後に書く形式で学生に指定している。縦書きでは、氏名の位置があまり下がり過ぎないように中央揃えにと指定しているが、横書きでは、右寄せでおか

しくないため右寄せに近い位置に指定している。

付録 1.7 7 の写真や絵を載せる場合への補足

写真や絵を載せる場合の留意事項を独立して記している。

付録 1.8 8 について

知っておいて欲しい事柄を補足している。学生、赤点、体育競技会など、間違わずに使って欲しい言葉の確認も促している。

付録 1.9 9 について

原稿についての最後の確認と、原稿の提出方法について記している。原稿収集後の取り扱いは出席番号順が好都合である。出席番号順で集まるように、学生に協力してもらっている。

病欠等で遅れた場合については、区別できるように、知らせてもらうことにしている。

付録 2 カレンダーについて

ここは、2.5.2.6 への補足である。

カレンダーを付けることがある (遅くても平成 10 年度から。平成 15 年度に 120 年分から 140 年分にしたのは、この宇宙の年齢の約 1 億分の 1 などより)。

期間は、平成 15 年度からクラスの学生の誕生日と百歳の誕生日と保護者の一部の誕生日も入る期間で、近年は 1971~2110 年である。約 5 万日の文字内容は自作のコンピュータプログラムで生成している。

文字サイズは 2 ポイントである (当初は 1.8 ポイント)。印刷機でこのサイズの文字を十分再現できなかったため、人数分をレーザプリンタで出力している (カレンダーの 1 ページ目は、情報関連等記事を印刷した用紙の左半分)。

付録 3 懸案の事項等

保留されている問題や懸案の事柄を順不同に挙げる。

○文集を作成すること自体について

以前より、災害時や成人式や同窓会に備えた情報の持ち方も難しくなっている時世である。将来、文集のようなものを残すことや残したことが否定される時代が来るかどうかは分からない。

○完成媒体について

ミニ文集は初期から印刷媒体でのみ作成している。

前世紀の内から、文集に限らず光ディスクなどに作り込む例を聞くことがある。その場合、画像や音声や動画像なども含められる。情報通信ネットワークの時代であるとか高等専門学校ならばという気持ちもなくはないが、印刷媒体は多少かさばりはするものの他の手段を要さず直ぐに読めること、時代と共に容易さは著しく変化しているものの他への転載がしやすい電子データそのものではないこと、今ある他の手段は媒体や用いている技術が手軽に何十年も同じく使えるという保証がないことの辺りを気にしている。

先々も読めるためにはミニ文集の所持か存在が前提である。約40倍という冗長性から、数十年後もクラス毎に何冊かは残っていると期待している。

○画像の挿入

画像の挿入も試して欲しくて、携帯電話が普及する前頃、フロッピーディスク経由でパソコンにデータが移せるデジタルカメラを授業中に貸し出した時期があった。その頃、利用して1人や複数人で写った写真などを載せた学生も少なからずいた。

携帯電話から取り出した記録媒体をパソコンから読み取るリーダーを貸し出した時期もあった。

今ではスマートフォンが普及して写真は取りやすくなったが、ワープロに移すのに少し前まで学生が一律に使える手段はなかった。そのため、写真などを原稿に挿入する学生は限られていた。

令和2年度から本校は遠隔授業システムが導入されており、これはスマートフォン内のデータとパソコンのワープロを容易につなぐ。しかし、令和2年度は学生に案内する余裕が見つけられなかった。

○原稿内での項番と見出しの記載

提出物によくある項番と見出しを付けるのを必須事項にすると、ひとつの練習機会にはなるものの文集らしくなくなると思い、必須とはしていない。学生が判断して付けることは自由としている。

○製本日等の記入欄

ミニ文集のどこかに学生が氏名や記念に製本年月日を書き入れる欄を設けるかどうか思案することがあるが、これまで設けていない。

○ホチキスの針の錆びないもの

現在使用しているホチキスの針は、メッキされているものの鉄製である。ホチキス止めで数十年間保管された資料で状態が良いものもある一方で、針の錆びで汚しているものを見かける場合もある。何か

対策したいが、現在の厚みに必要な針の長さ(1 cm)について錆びない針はまだ探し当てていない。

○両面印刷の活用

製本用印刷はB4判用紙に片面である(2.5.1)。もし両面印刷にして、折った山の側で綴じたときの弊害は、その反対側の縁でページ毎の凸凹ができやすいことである。そこを切り揃えられれば解決し、用紙を節約でき、ホチキスや針はよく使われる小型のもので済み、防錆の針も使える。しかし、今のところ、安全で廉価な道具はまだ探し当てていない。

○綴じ後の針が盛り上がりがないホチキス

初期から綴じ後の針に盛り上がりができるホチキスを用いていて、簡易製本に平口で盛り上がりもなく手順も含めている(2.7.1)。見方を変えれば、盛り上がりができる場合の処置も、経験機会と考えることはできる。

盛り上がりがないと称しているタイプもあるが、最近試した一機種は完全ではなかった。

○綴じ後の針の盛り上がりもなく適切な手段

現在、平口を用いている(2.7.1)。実際にはホチキスの針に力が若干斜めに加わっており斜めにややつぶれやすい。盛り上がりを実際に真っすぐつぶせるような簡易な器具はまだ探し当てていない。

○簡易製本の器具の増量

簡易製本は割りと早くから2セットの器具で行っている(2.7.2)。作業時間の短縮のため、目が届くことも考慮したもう1セットの追加を考えている。

○原稿提出前のチェック

実施要領に沿っているかどうかの確認は各自の責任としているが、種々の確認漏れのケースが散見されるため、原稿提出前に予め指定したチェック役にも、実施要領に沿っているか確認してもらうこと(多くの場合に、2人ずつ相互にチェック役になる。確認しながら、別途配布のプリントに記載した各確認項目に印を付ける)としていた時期がある。効果が高いとは言えず、令和2年度から本人による確認だけに戻した(付録1.9)。

○カレンダーの部分をインクに

カレンダーの部分はレーザープリンタで出力している(付録2)。トナーはインクより劣化が早いのはとの懸念があるが、現状で筆者に他の選択肢はない。

1年 学科 番 氏名 ①

演習[実施要領兼自己チェック用紙] (操作例はMS Word2016。以下、大部分で文字10ポイント、行ピッチ12.5ポイント、1ポイント=0.35mm)

★★ 1 あらまし ★★

《フォントは主にUDデジタル教科書体N-Rとゴシック体太字を使用》

- ・クラス毎にミニ文集を作成する。最終的に皆が1冊ずつ自クラスのミニ文集を手にする。◆下書き前に①の3まで、入力前に②の8まで、提出の前週に9まで、目を通すこと。
- ・◆各人の持ち分はA4判1ページ(超えた分は用いない)。
- ・◆ワープロ演習のひとつとして実施する。◆実施要領①と②は全て学習内容や課題である。加減事項もある。◆違反は減点。◆遅れや違反によってはミニ文集に載らない。◆現担任と来年度同学科2年担任へは差上げる予定である。

★★ 2 主な流れや日程 ★★ 《フォントは主に楷書体を使用》

- (1)◆宿題として後期中間試験後最初の授業前までに下書き(③で指示)を完成しておく。中間試験直後の休養のためにも、11月中頃までにある程度書いておいた方が無難。
- (2)授業時間は後期中間試験明けから30分程度を2回だけワープロ入力にあてる。1月は実力試験や突発の宿題やインフルエンザ感染等もあり得るので入力を早めに進めること。
- (3)◆原稿提出は1月で、M:13日、E:13日、C:18日、A:14日。
- (4)中村が印刷や道具の準備。製本は各自(右と同じ、別途連絡)。

★★ 3 執筆内容や字数について ★★ 《フォントは1と同様》

- ・下記(ア)~(ケ)の内容を順に書く。若干詩的に言うなら、日々新しいあなたの瞬間をギュッと凝縮した内容を期待する。
- ◆文集なので、まず、(ア)自由なテーマ・内容で、7百字以上で書く:例えば、最近思う事、語りた事、皆へ、今の自分、人生設計、将来貢献したい事、これらのために決意・抱負、XX歳の自分へ、未来日記X年X月X日。下記(イ)~(ケ)の話題例をテーマに選んでも可。◆執筆実況で字数稼ぎは減点
- ◆下記(イ)~(ケ)を順に、(オ)は2百字以上、他は百字以上
- (イ)自己紹介:例えば、出身校、生い立ち、趣味・嗜好(しこう)・お薦め・推し、長所、苦手、特技、部活等、名前の由来、好きな言葉、自分のOnly One、郷里の紹介・市外なら都城との比較
- ◆ここに限らず、誕生日、電話番号、詳細な住所や通学経路や帰省経路を書くのは不可。
- (ウ)高専を選んだ訳と今の学科を選んだ訳(もし第一志望でなかった場合は志望のひとつにした訳を):主な訳やささいな訳
- (エ)これまで、または最近のこと:例えば、今まで出会ったすごい人、今までのでき事、頑張った事、継続中の事、マイグム、高専の知名度、XXから聞いた高専のうわさと真相、高専や学科の長所・短所・発見・後悔・希望、他校と比べて、課外活動、寮、お薦めの物・本・場所・Webサイト・アプリ、試験対策・取り組み、休日の過ごし方、マイストロード 諸々、自分の成長。
- (オ)将来就きたい仕事:就きたい仕事として関心を持っている職種か分野か企業など(自分の考えを。今の考えでよい。先々変化して構わない。候補を複数書いてもよい。)と、理由も書く。更に、その仕事での夢、研究・開発・製作・探求・残したい物事の中からひとつ以上書く。必要な勉強・資格・進路などで今知っている事がもしあれば、それにもふれる。
- (カ)これからの自分について5つ以上(夢、計画、希望、野望、理想等):例えば、今したい事・すべき事、今年・在学中・卒業後・数十年後等の目標・夢・行きたい所・したい事・趣味、XX年後の仕事・社会・世界・家庭と自分、夢の簡条書き。
- (キ)将来の自分へのメッセージ(希望や期待も可)
- (ク)将来の自分の子供か子の世代へのメッセージ(希望、期待も可。性別等想定してもしなくても双方述べても可)
- (ケ)今後50年以内に起こる事柄・変化・発明の予言・予想を7つ以上(◆自分のこと以外を。◆何かが終るやバージョンが進むという類は減点。◆ここに限らず、元号や各国等の政権関連は書くのは不可(あなたや皆の安全のため。))
- ・どこでも簡条書き、太字、下線、傍点、写真、図、縁取りも可。

- ・◆(ア)の内容は最初に題を書く。◆(ア)の内容は、必ず、複数段落で書き、各段落の頭は字下げする。◆(イ)以降に見出し等は無くても可(例:先頭の字下げや、先頭に何か記号や、空行。付ける例:太字や下線付きの見出し)。◆見出し中や見出し代わりに、ア~ケの字を使わない(あれば減点)。
- ◆(ア)と(イ)の内容の切り替わりが分かるようにする(見出しを付けない場合でも、書き出しの工夫や空行で可能)。
- ◆文、写真等々は各自のオリジナルなものに限る。◆それらに現れる人物等から文集に載せる承諾を得てないのは不可。◆他の人の暴露(ばくろ)話は不可。◆文も他も、他人の著作物(Webのもの)の流用、添削物は不可だが、自分が入った写真と引用(誰がどんな題で何の何ページにいつ公表したか情報を明記し、少量の文をそのまま載せること)は可。
- ◆不愉快になる人・団体・国民・政権・民族や、傷つくる人のいる可能性のある内容は不可。◆伏せ字、イニシャル、「某」などを使った表現、下品な内容、デマ、嘘は不可。◆内容の責任は各自が負うものとする。私的な事柄の掲載は各自で判断。
- ・もしも他に発表等や文集の計画がある場合は、そちらを主と考え、このミニ文集に内容の重複があって構わない。
- ・次の場合は委員長を通じて中村と打ち合わせること(連絡だけでも12月16日までに):表紙等や集合写真や寄せ書き(手書きや読み取ったものも可)の印刷用原稿を作りたい。なお、中村が表紙を作るとき、内容は、年度・クラス名、製本年月、課題名程度の予定。

- ◆次に該当し更に①と②が無違反の時、原稿提出に(次の★印毎に)3割程度(軽微な課題1回分相当、但し、全課題評価の満点まで)加点する:★下書き(③で指示)を11月15日12:45までに見せに来た、★字数が2100字(または、1100ワード)以上で1月の原稿提出日より前の日に②の9を済ませた実施要領①と②と原稿を提出し同時に口頭で字数(または、ワード数)を申し出た、★縦書き(5を参照)を700字以上含む、★英文を350ワード以上含む(英語のミスは減点しないが、Web翻訳のままは減点)、★写真や自作のイラスト(図形機能やペイントやドローソフトで作成か、手書きを撮影)を含む(7も参照)、クラス全員が了解済みの次のページの印刷用原稿を作った(共同作業なら加点は人数割プラス0.5割):★表紙、★集合写真、★寄せ書き(手書き可)。

《このあとはははらく、種々のフォントを使用》

★★ 4 原稿作成に関する設定・操作について ★★

- ◆余白を「レイアウト」「ページ設定」「余白」等で設定。◆印刷内容の上端は用紙の上から18mmの所、左端は用紙の左から30mmの所に、右端は右から15mmの所にあること。下端は下から18mm以上のこと。◆これらは、±4mm以内の誤差は減点しない。更にはみ出た部分は載らないかも。
- ◆実は、上記の余白設定が効かない機能もあるため、各自が、印刷してから上下左右の余白を実測して、規定外なら、はみ出ないように調整すること。予め、印刷プレビューで、やや拡大表示して、画面で(画面を傷つけないように)測ると印刷の失敗を減らせる。◆用紙の向きは、「余白」の所で、原則「縦」を確認。横長も可だが、製本用印刷では縦長で使う。
- ◆半分程度以上を段組み(2段以上)にすること(段組みにしたい範囲をドラッグして「レイアウト」「段組み」で種類等を選べばよい(このプリント①は(ツグー)の下の最初だけ段組み無し、つまり1段で、他は2段)。段の「間隔」は自由(例えば、1字~2字分)。氏名番号等(後述)は1段で記載すること。
- ◆行数や1行あたりの字数は自由(例:「レイアウト」「ページ設定」)。
- ◆行間隔(正確には行ピッチだが、Wordではこう呼ぶ)を変えたいなら対象範囲をドラッグして、「ホーム」「段落」の中央の「行間」を「固定値」にし、「間隔」を15ポイント

1年

学科 番 氏名

②

前後で選択または手入力(小数も可)。◆10ポイント未満の字を、ルビ(後述)以外で、中村の許可なく使ったら減点。
 ◆1ページを超したとき、まず行数や行間隔を変えてみる。更には、文字サイズを10ポイントにし行間隔の調整や、余白を先述の減点限度までせまく調整。最後の余白が広過ぎるとき、せばめるのも、考え方は同じ。◆大部分の行間の白いすき間が、文字サイズと同程度以上の原稿は減点。
 ・この課題は、指示した以外は、フォント・文字サイズ・スタイルをどこでも自由を選んでよい。変更は容易なので色々試して気に入ったのを使ってよい。手書きフォントも可。
 ・◆ページ番号やヘッダー、フッターの機能は用いないこと。

★★ 5 縦書きにする場合への補足 ★★

・縦書き・横書きはいつでも変更可。◆全体を縦書きにする例:「レイアウト」「ページ設定」「字数と行数」で「縦書き」。
 ◆この課題で、英単語等は字が寝ていて構わない。◆瘦た字を縦にする例:範囲指定して「ホム」「A」「縦中横」。テキストボックスを使うと、縦書きや横書きの文がどこでも置ける(参考:「挿入」「テキストボックス」や「図形の書式設定」「線」「線なし」)。

★★ 6 氏名番号等の入れ方 ★最も注文の多い1行か1列(「;」)

・◆氏名番号等は、次の形式で、全体の前側に1段で記載:
 姓 名 (ローマ字で姓名) 何学科 1年 何番
 ・◆字は、丸カッコとローマ字は半角で、他は全角。◆日本語での姓と名に別々にルビ(読みがな)をふる(ルビは「ホーム」で「ア」のボタン。付けたはずのルビが表示されないうとき、その行だけ行間隔を16ポイント前後にしてみる。行間隔の変更は、4を参照)。◆ローマ字(縦書きの場合、字は寝かす)は姓名の順で(こだわりがあるなら姓名も可)。◆姓は全て大文字、名は頭文字だけ大文字で。◆「(」の前と、ローマ字の姓と名の間に、半角スペースを1字か2字入れる。
 ◆更に学科名の前に全角スペースを2字入れる。◆学科名は略さない正式名称で。◆学年と出席番号の数字は全角の等幅フォント(大概は、全角のフォント名にPがないもの)の算用数字(縦書きの場合は漢数字)で。◆日本語の姓と名の間と、「1年」の前後に、全角スペースを1字入れる。◆横書きなら、「番」が右端から(本文の字の位置で)ほぼ3字目になるよう、行全体をうまくずらすこと。◆縦書きなら、氏名番号等の列は上下中央揃え(「ホーム」「III」)に。
 ・◆段組例:最初に段組設定なし(=1段)で全て入力し、後で氏名番号等以外を2段組にしてもよいし、最初に2段組にして本文だけ入力し、後で以下の操作で氏名番号等の行を付けてもよい:冒頭に、先述の氏名番号等の字を少し入力し、それを範囲指定し、「レイアウト」「段組み」「1段」を選ぶ。
 ・◆もし本文を英語で書いたとき、氏名番号等も英語でもよい。◆もしそうしたときだけは、数字には半角の算用数字を使い、丸カッコ内には姓名をルビ付きの日本語で書く。

★★ 7 写真や絵を載せる場合への補足 ★★

・◆自分が入った写真以外は、自分で撮影や描いたものに限る(スマホで撮り工夫して経由すれば Word で使用可)。複数人が写った写真は皆で使用可。◆写真や絵は、写っている人全員や、写ったもののデザインの権利者や会社の承諾がないのは、後から分かってても不可(合法的な写り込みは可)。
 ・◆今回製本用印刷をする印刷機は、このプリントと同様に、にじみやかすれが生じる。画像があると、一緒に印刷する原稿2枚で文字の太さが若干変わる。灰色の再現は不得意。
 ◆写真の大きさは任せるが、顔は、小さいと違う表情で印刷され易い。直径2cm以上が割と無難(保証はできない)。

★★ 8 色々(ここも重要) ★★《主に明朝体とゴシック体太字》

・文書全体を範囲指定したいときは、コントロールA、解除はCtrl/C。
 ・◆ワープロ画面の赤や緑の波線や青の二重線は、印刷はさ

れないが、誤記等の可能性の警告なので、一度は字や文を確認すること。クリックして右ボタンで、一応理由が出る。
 ・◆内容の完成時や印刷前に「校閲」「スペルチェックと文章校正」をすること。注意:◆それでも、言い回しや漢字の間違ひはソフトには分からないことが多いので、自分がしっかりすること。
 ・◆誤字脱字や不自然な空白や改行等もないよう、自分がしっかりする(製本後は皆の手に残る)。◆次の事で誤解のある原稿は減点:皆さんは生徒ではない、60点丁度は赤字ではない、本校に体育祭はない、高専祭と文化祭は異なる。
 ・◆引用部分やアスキーアートは字数から引くこと。◆文字カウント機能(「校閲」「文字カウント」)は、範囲指定部分にも働き、単語数(=ワード数)も表示される。◆Wordの編集画面左下に表示されている字数も参考になる。皆さんは約1600字以上だけど、①と②でな一んと約6400字(「;」)削れない訳は、削ると規定外原稿や質問が必ず増すから。(脱線 ↑ チコちゃんをまねて言うてみてください。)

・◆できあがりや総ページ数(この課題は1ページ)は、印刷プレビュー(「ファイル」「印刷」)などで直ぐに確認できる。
 ・◆左揃え(≡)だと、通常、文の句読点が行頭に来そうなとき、直前の文字も行頭に移動。これを避けたいとき、その行などで「ホーム」で「≡(両端揃え)」を選ぶと改善することが多い。
 ・◆もし段組みの直後の2ページ目の先頭の空行を除去できないとき、一旦1段組みにして、無駄な空行を除去してから段組みにしてみる。段組みが壊れたら、中村へ連絡を。
 ・◆提出原稿は、黒以外の使用(灰色は画像のみ可)や、筆記具による書き込みは、不可。◆紙などの貼り付けや、裏に印刷や書き込みがある場合も、不可。◆原稿は折らないこと。
 ・◆フォントや絵や句読点の位置によってはうまく印刷されないことがある。設定やフォントや内容やプリンタの変更を試す。自宅のインクジェットプリンタで、普通紙でにじむ場合は、用紙を変えるか、コンビニ印刷(別途連絡)を活用する。
 ・中村が、右とじ(似た例:国語の教科書)用の製本用印刷(提出原稿を2人分ずつ87%に縮小しB4判白黒印刷)と、ホチキス(ステプラー)や製本テープ等の準備をする。各自で、印刷物の拾集、折り曲げ、ホチキス留め、製本テープ貼り(別途連絡)。
 ・◆念のため、3月まで原稿のファイルは削除しないこと。
 ・◆印刷をコピーとか、プリンタをコピー機とか言わないように。◆プリンタ出力は印刷であり、コピーではない。
 ・◆やむなく半角カタカナを多用した。普通は用いない方がよい。
 ◆この課題では構わないが、算用数字と英字は半角文字使用が常識の場合も多いのは覚えておくこと。◆ミニ文集は著作物で、更に私的な事柄も含むので、皆が取扱いに注意。

★★ 9 原稿提出の前週に確認 ★★ 《7月分は1と同様》

・◆原稿提出前々日までに、原稿を、実施要領①と②の、個々のひし形印の内容で最後の確認をし、完了(または、内容をしっかり把握)したとき、そのひし形印を丸で囲んでおくこと。◆不備は、修正や完了してから丸を付けること。◆確認は執筆者の責任。◆丸付けのめれや無記名の提出は不可。
 ・◆提出は、①と②とA4判原稿1ページを、とじないでこの順番で重ねて用意し、各自が提出日に、定期試験での自分の列の先頭の学生に持参して渡すこと(教室授業であっても)(試験時列先頭の学生は、集めに回らないこと)。◆必ず情報基礎授業の始業前までに渡す。でない遅れ扱い。
 ◆試験時列先頭の学生は、3枚重ねたまま出席番号順にしながら受け取り、授業開始直後に中村に提出。◆もし遅れても、きちんとしたものを提出(中村に手渡して)。◆遅れた学生は、①の上側に理由を赤で具体的に書いておく(締め切り前日等の種々の機器等の不調等は、同情はするが遅れの正当な理由としては加味できない)。◆①と②は記者名には返却するが、再提出は受理しない。原稿は返却しない。

付図1 実施要領 (2)2ページ目

分散登校および遠隔授業による新型コロナウイルス感染症に対する 学内感染リスク低減効果の統計的評価方法と解析結果

田中 守¹

Statistical evaluation method and analysis results of the effectiveness of distributed school attendance and remote learning on reducing a risk of infection in a school

TANAKA Mamoru¹

(Accepted September 22, 2021)

Abstract As a countermeasure against SARS-CoV-2, which is still raging in the world as of September 2021, our school has been implementing distributed school attendance and remote learning. We were interested in how much the risk of infection in the school could be reduced by distributed school attendance and remote classes. However, since the actual number of infected students at National Institute of Technology (KOSEN), Miyakonojo College is small, it is difficult to evaluate the risk of infection on campus using this data. In this paper, we give a model of a risk of infection in our school, and report the results of the effect of distributed school attendance and remote learning on reducing the risk of infection in the school.

Keywords [SARS-CoV-2, A risk of infection in the school, Distributed school attendance, Remote learning]

1 序論

2021年9月現在も世界で猛威を振るっている新型コロナウイルス感染症の対策として、これまで本校では分散登校および遠隔授業が実施されてきた。そこで、分散登校および遠隔授業により学内の感染リスクがどれだけ低減できたのかについて興味を持った。しかし、実際の都城高専の感染者数は少ないため、その感染者数を用いた学内感染リスクの評価は難しい。本稿では、1つの学内感染リスクのモデルを与え、分散登校および遠隔授業による学内感染リスクの低減効果を計算した結果の報告である。

より具体的には、通生・寮生および居住地・帰省

先別の学生数、宮崎県および鹿児島県の年齢別人口統計データ、実際の宮崎県および鹿児島県が毎日HPにて発表するデータ（日付、年代、地域）等を用いて、登校できる学生の中の新型コロナウイルス感染症にかかっている学生数の確率的な平均値を学内感染リスクとして与えた。さらに、分散登校および遠隔授業を行った場合と行わなかった場合について、その差を低減効果として与えた。ただし、登校できる学生でも症状があれば登校しないことや、新型コロナウイルス感染症にかかっているとしても他の学生に感染させるとは限らないなど、より正確な学内感染リスクはさらに小さくなるものと考えられる。

2 学内感染リスクのモデル

ここでは、まず小中高校における感染者数から10代における年齢別の感染者数を推定する。次に、学生を通生、寮生および都城市・三股町、それ以外の宮崎県内、鹿児島県、専攻科生に分け、県の市町村の年代別人口データを用いて、各地区における本校の学生の割合を与え、それを用いて学内感染リスクのモデルを与える。また、以下の議論では、有効数字は高々1桁であるとして考察する。

2.1 都城高専および市町村の年齢別人口

ここでは計算の単純化のために、宮崎県・鹿児島県以外の寮生、鹿児島県の通生はいないものとし、専攻科生はすべて都城市・三股町の通生とし、各地域の通生・寮生数については、下記の表のように仮定する。

表1 都城高専学生数 (単位:人)

都城市・三股町の本科生の通生数	$C_{mm} = 450$
都城市・三股町以外の宮崎県の通生数	$C_{mz} = 50$
都城市・三股町以外の宮崎県の寮生数	$D_{mz} = 220$
鹿児島県の寮生数	$D_{kg} = 80$
都城市・三股町の専攻科生の通生数	$C_{ad} = 50$

下記の表の宮崎県の市町村の年齢別の人口数は、令和2年:年齢別人口構成表(令和2年10月1日現在)における年齢別人口構成(市町村別)¹⁾を用いた。また、鹿児島県の年齢別の人口数は、年齢別人口統計表一覧(令和元年)²⁾を用いた。

表2 各地区の年齢・年代別人口 (単位:人)

都城市・三股町の15~19才	$H_{1mm} = 8823$
都城市・三股町以外の宮崎県の15~19才	$H_{1mz} = 40334$
鹿児島県の15~19才	$H_{1kg} = 74206$
都城市・三股町の20代	$H_{2mm} = 14177$

2.2 10代中の年齢別の感染者数の推定

文部科学省の「学校における新型コロナウイルス感染症に関する衛生管理マニュアル~「学校の新しい生活様式」~(2021.4.28 Ver.6)(令和3年5月28日一部修正)³⁾により、令和2年6月1日~令和3年4月15日までの小学生から高校生までの感染者数は、小学生6183人、中学生4072人、高校生7046人である。これより、年齢とともに指数的に感染者数が増加していることが推測できる。そこで、年齢 x に対して、その感染者数 y を $y = e^{ax+b} +$

c と仮定して、

$$\int_6^{12} (e^{ax+b} + c) dx = 6183$$

$$\int_{12}^{15} (e^{ax+b} + c) dx = 4072$$

$$\int_{15}^{18} (e^{ax+b} + c) dx = 7046$$

を解くことで、 a, b, c を求めると

$$y \approx e^{0.415x+0.332} + 956$$

となる。これを用いると、10代の感染者において、高専生と同じ学年の学生が占める割合は約70%であり、特に低学年生の占める割合は約30%となる。

表3 10代の感染者中の各学年が占める割合

高専生と同じ学年	$R_{ks} = 0.7$
低学年生と同じ学年	$R_{ksl} = 0.3$
高学年生と同じ学年	$R_{ksh} = 0.4$

このことから、少なくとも令和3年4月15日までは、低学年のみの分散登校にすることで、感染のリスクを半減させることができたことが推定できる。本稿では令和3年度前期までにおいて同様の傾向が続いたと仮定して、この感染者の学年別の割合を用いる。

2.3 感染者が登校できる都城高専の学生である確率

分散登校は低学年のみ登校する分散登校を意味し、各期間における各地区、各年代の感染者1人が登校できる都城高専の学生である確率 p を、それぞれの場合に対応して下付きの添え字をつけて、下記のように設定する。(ここでいう「登校できる」とは、例えば、分散登校では低学年が全員登校でき、高学年生は全員登校できないという意味であり、風邪症状等があっても登校できると仮定している。)ただし、入寮で居住地が変わった場合も想定し、その1週間後まで前の居住地の感染者データを用いる期間も設ける。

2.3.1 全員登校できる期間の内、入寮日から1週間以内

(1) 都城市・三股町の10代

$$p_{mm1} = \frac{C_{mm}}{H_{1mm}} \times R_{ks}$$

(2) 都城市・三股町以外の宮崎県の10代

$$p_{mz1} = \frac{C_{mz} + D_{mz}}{H_{1mz}} \times R_{ks}$$

(3) 鹿児島県の10代

$$p_{kg1} = \frac{D_{kg}}{H_{1kg}} \times R_{ks}$$

(4) 都城市・三股町の20代

$$p_{ad1} = \frac{C_{ad}}{H_{2mm}}$$

2.3.2 全員登校できる期間の内、入寮日から8日目以降

(1) 都城市・三股町の10代

$$p_{mm2} = \frac{C_{mm} + D_{mz} + D_{kg}}{H_{1mm} + D_{mz} + D_{kg}} \times R_{ks}$$

(2) 都城市・三股町以外の宮崎県の10代

$$p_{mz2} = \frac{C_{mz}}{H_{1mz} - D_{mz}} \times R_{ks}$$

(3) 鹿児島県の10代

$$p_{kg2} = 0$$

(4) 都城市・三股町の20代

$$p_{ad2} = \frac{C_{ad}}{H_{2mm}}$$

2.3.3 分散登校入寮日から1週間以内

(1) 都城市・三股町の10代

$$p_{mm3} = \frac{C_{mm}}{H_{1mm}} \times R_{ksl}$$

(2) 都城市・三股町以外の宮崎県の10代

$$p_{mz3} = \frac{C_{mz} + D_{mz}}{H_{1mz}} \times R_{ksl}$$

(3) 鹿児島県の10代

$$p_{kg3} = \frac{D_{kg}}{H_{1kg}} \times R_{ksl}$$

(4) 都城市・三股町の20代

$$p_{ad3} = 0$$

2.3.4 分散登校入寮日から8日目以降

(1) 都城市・三股町の10代

$$p_{mm4} = \frac{C_{mm} + D_{mz} + D_{kg}}{H_{1mm} + D_{mz} + D_{kg}} \times R_{ksl}$$

(2) 都城市・三股町以外の宮崎県の10代

$$p_{mz4} = \frac{C_{mz}}{H_{1mz} - D_{mz}} \times R_{ksl}$$

(3) 鹿児島県の10代

$$p_{kg4} = 0$$

(4) 都城市・三股町の20代

$$p_{ad4} = 0$$

2.3.5 分散登校終了後1週間以内（高学年が入寮後1週間以内）の場合

(1) 都城市・三股町の10代

$$p_{mm5} = \frac{C_{mm} + D_{mz} + D_{kg}}{H_{1mm} + D_{mz} + D_{kg}} \times R_{ksl} + \frac{C_{mm}}{H_{1mm}} \times R_{ksh}$$

(2) 都城市・三股町以外の宮崎県の10代

$$p_{mz5} = \frac{C_{mz}}{H_{1mz} - D_{mz}} \times R_{ksl} + \frac{C_{mz} + D_{mz}}{H_{1mz}} \times R_{ksh}$$

(3) 鹿児島県の10代

$$p_{kg5} = \frac{D_{kg}}{H_{1kg}} \times R_{ksh}$$

(4) 都城市・三股町の20代

$$p_{ad5} = \frac{C_{ad}}{H_{2mm}}$$

2.3.6 遠隔授業のみの場合

いずれの場合も $p_{mm6} = p_{mz6} = p_{kg6} = p_{ad6} = 0$ とする。

2.4 学内感染リスクのモデル

指定した期間（例えば令和2年度前期など）において、全員登校できる期間の内、入寮日から1週間以内における都城市・三股町の10代の感染者が n_{mm1} 人いた場合に、その中で登校できる都城高専の学生の人数を $n_{mm1}p_{mm1}$ 人と推定する。同様に、入寮期間や分散登校に関する他の期間、地区、年代に対しても、同様に人数を推定する。このとき、指定した期間における学内感染リスク[単位：人]を

$$\lambda_{\text{指定した期間}} = n_{mm1}p_{mm1} + \dots + n_{ad6}p_{ad6}$$

とする。これは、指定した期間において登校できる都城高専の学生中の感染者の人数を推定するものである。

また、全員登校できる期間の内、入寮日から1週間以内における都城市・三股町の10代の感染者 n_{mm1} 人に対して、その中の登校できる都城高専の学生の人数を確率変数 X_{mm1} 人で表すとする。登校できる学生の中に感染者がいた場合でも他の学生が感染しやすくなることはないかと仮定すると、この確率変数 X_{mm1} は、二項分布 $B(n_{mm1}, p_{mm1})$ に従う。同様に、入寮日や分散登校に関する他の期間、地区、年代に対しても確率変数が定義でき、二項分布に従う。このとき、

$$X = X_{mm1} + \dots + X_{ad6}$$

の平均は

$$E(X) = \lambda_{\text{指定した期間}}$$

となる。ここで、考えている二項分布 $B(n, p)$ は n が大きく、 p が小さいとして、近似的にポアソン分布 $P_o(np)$ とみなせるとする。このとき、ポアソン分布は再生性を持つため、確率変数 X は平均 $\lambda_{\text{指定した期間}}$ を持つポアソン分布 $P_o(\lambda_{\text{指定した期間}})$ に従う。

3 学内感染リスクの低減効果

宮崎県の「宮崎県における感染者発生状況一覧」HP⁴⁾、鹿児島県の「県内・国内外の発生状況」HP⁵⁾の公表データを用いて、令和2年度前期、令和2年度後期、令和3年度前期の休業期間以外の期間について、学内感染リスクの低減効果を与えたい。ここで開始・終了日については、開寮日、閉寮日を基準とした。また、宮崎県および鹿児島県のHPデータから公表日の前日を感染判明日とした。

本校における分散登校と遠隔授業を行う大まかな基準は、都城市・三股町の直近1週間の新規感染者数が20人（人口10万人あたり約10.8人）を超えた場合に分散登校、都城市・三股町が県の指定する赤圏域になった場合に遠隔授業である。

3.1 令和2年度前期

令和2年度前期において例年通り、5月の連休（準閉寮）を除き4月4日～5月2日、5月6日～8月12日に対面授業を行ったと仮定する。このとき、4月4日～4月10日および5月6日～5月12日が「全員登校できる期間の内、入寮日から1週間以内」であり、4月11日～5月2日および5月13日～8月12日が「全員登校できる期間の内、入寮日から8日目以降」である。この「全員登校できる期間の内、入寮日から1週間以内」における都城市・三股町の10代の感染者は $n_{mm1} = 0$ 人、都城市・三股町以外の宮崎県の10代の感染者は $n_{mz1} = 3$ 人、鹿児島県の10代の感染者は $n_{kg1} = 1$ 人、都城市・三股町の20代の感染者は $n_{ad1} = 0$ 人であり、

「全員登校できる期間の内、入寮日から8日目以降」における都城市・三股町の10代の感染者は $n_{mm2} = 0$ 人、都城市・三股町以外の宮崎県の10代の感染者は $n_{mz2} = 17$ 人、鹿児島県の10代の感染者は $n_{kg2} = 16$ 人、都城市・三股町の20代の感染者は $n_{ad2} = 2$ 人である。分散登校や遠隔授業の期間はないと仮定しているため、これを用いた学内感染リスクは

$$\lambda_{R2 \text{ 前期}}(\text{例年通り}) = n_{mm1}p_{mm1} + \dots + n_{mm2}p_{mm2}$$

であり、実際に計算すると

$$\lambda_{R2 \text{ 前期}}(\text{例年通り}) = 0.04 \quad \text{人}$$

であった。さらに、感染者数が0人である確率をポアソン分布を用いて計算すると、約96%であった。

一方で、実際の学校運営は、5月10日まで休業とし、5月11日より遠隔授業および分散登校（授業内容は遠隔授業を学校で受けるものであり、2学年ず

つそれぞれ2週間のみ、登校しないで遠隔授業を受けてもよい）が行われた。すべて遠隔授業として登校していない状況と同等とみなすと、学内感染リスクは

$$\lambda_{R2 \text{ 前期}}(\text{分散・遠隔あり}) = 0 \quad \text{人}$$

である。いずれにしても、学内感染リスクはほぼないとみなせるため、登校しないことによる学内感染リスクの低減効果もほぼないと考える。

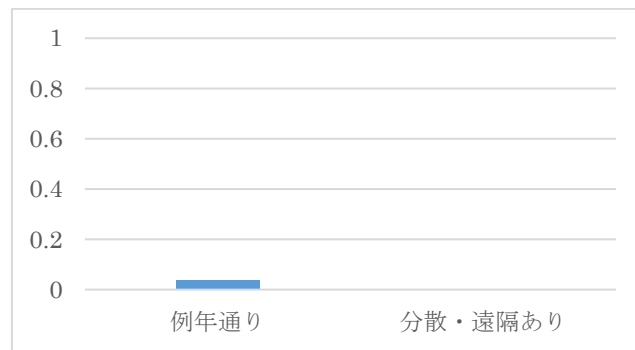


図1 学内感染リスク (令和2年度前期)

令和2年度前期において、実際の学生の感染も（休業日を含めて）発生していない。

3.2 令和2年度後期

令和2年度後期において例年通り、冬季休業を除き9月27日～12月26日、1月5日～2月27日に対面授業を行ったと仮定した場合も前期と同様に計算すると、学内感染リスクは

$$\lambda_{R2 \text{ 後期}}(\text{例年通り}) = 0.87 \quad \text{人}$$

であった。さらに、感染者数が0人である確率は約42%、1人以下である確率は約78%、2人以下である確率は約94%、3人以下である確率は約99%であった。

一方で、実際の学校運営は、9月27日～12月26日まで対面授業、冬季休業を挟み、1月5日～2月27日まで遠隔授業が行われた。したがって、9月27日～10月3日が「全員登校できる期間の内、入寮日から1週間以内」で、10月4日～12月26日が「全員登校できる期間の内、入寮日から8日目以降」であり、1月5日～2月27日が「遠隔授業のみ」である。このとき、学内感染リスクを同様に計算すると

$$\lambda_{R2 \text{ 後期}}(\text{分散・遠隔あり}) = 0.14 \quad \text{人}$$

であり、感染者数が0人である確率は約87%、1人以下である確率は約99%であった。

したがって、令和2年後期の学内感染リスクの低

減効果は

$$\lambda_{R2 \text{ 後期}}(\text{例年通り}) - \lambda_{R2 \text{ 後期}}(\text{分散・遠隔あり}) = 0.73 \text{ 人}$$

である。これは、登校できる都城高専の学生の感染者を0.7人程度減らしたと見積もることができる。

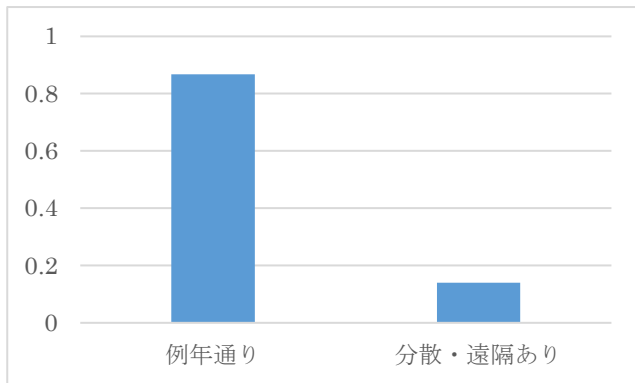


図2 学内感染リスク (令和2年度後期)

令和2年度後期において、登校日には実際の感染者は発生していない。しかし、登校日の2日後の12月28日に2人、登校日の3日後の12月29日に1人の学生の感染者が確認されている。これについて、感染者が1人でも確認されたことは確率的にまれなケースといえるが、その1人から感染が広がるとすると3人であることについて多いとは言えない。

3.3 令和3年度前期

令和3年度前期において例年通り、5月の連休(準閉寮)を除き、4月3日～5月1日、5月5日～8月12日に対面授業を行ったと仮定した場合、学内感染リスクを同様に計算すると

$$\lambda_{R3 \text{ 前期}}(\text{例年通り}) = 2.97 \text{ 人}$$

であった。さらに、感染者数が0人である確率は約5%、1人以下である確率は約20%、2人以下である確率は約43%、3人以下である確率は約65%、4人以下である確率は約82%、5人以下である確率は約92%、6人以下である確率は約97%であった。

一方で、実際の学校運営は、4月3日～4月24日まで対面授業、4月25日～5月1日まで分散登校、5月の連休を挟み、5月5日～5月8日まで分散登校による対面授業、5月9日～5月29日まで遠隔授業、5月30日～6月5日まで分散登校による対面授業、6月6日～8月12日まで対面授業が行われた。このとき、学内感染リスクを同様に計算すると

$$\lambda_{R3 \text{ 前期}}(\text{分散・遠隔あり}) = 1.67 \text{ 人}$$

であり、感染者数が0人である確率は約19%、1人

以下である確率は約50%、2人以下である確率は約77%、3人以下である確率は約91%、4人以下である確率は約97%であった。

したがって、令和3年度前期の学内感染リスクの低減効果は

$$\lambda_{R3 \text{ 前期}}(\text{例年通り}) - \lambda_{R3 \text{ 前期}}(\text{分散・遠隔あり}) = 1.3 \text{ 人}$$

である。これは、登校できる都城高専の学生の感染者を1人程度減らしたと見積もることができる。

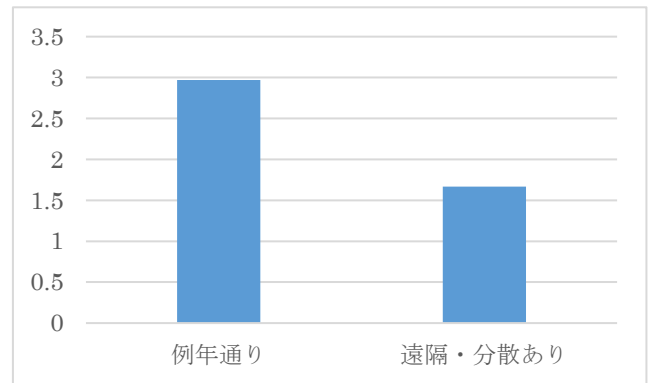


図3 学内感染リスク (令和3年度前期)

令和3年度前期において、遠隔授業中で、例年であれば登校日である5月19日に1人の学生の感染者が確認されており、登校できる都城高専の学生の感染者を実際に減らすことはできている。

また、実際の対応における学内感染リスクの大半は、4月25日～5月1日までの分散登校および8月6日～8月12日までの対面授業における学内感染リスクであった。この期間に感染者が出なかったことが幸いしたと考えられる。

4 まとめ

本稿では、本校の通生・寮生および居住地・帰省先別の学生数、宮崎県および鹿児島県の年齢別人口統計データ、実際の宮崎県および鹿児島県が毎日HPにて発表するデータ(日付、年代、地域)等を用いて、指定した期間における登校できる本校の学生の感染者数の確率的な平均値として学内感染リスクを与えた。その結果をまとめると、

(1) 令和2年度前期においては10代、20代の感染者が少なかったため、例年通りの対応でも学内感染リスクが低い状況であった。

(2) 令和2年度後期においては、学内感染リスクを0.87人から0.14人に、約0.7人分減らすことができた。

(3) 令和3年度前期においては、学内感染リスク

を 2.97 人から 1.67 人に、約 1 人分減らすことができた。

これらの低減効果を大きいとみるか小さいとみるかについては、学校運営において、「1 人も感染者を登校させたくない」とするか、「感染者が 1 人いたとしても他の学生に広げないような感染対策をとっている」とするかなどによって変わってくる相対的な問題であり、本稿で結論を出すことはできない。しかし、学内感染リスクについて数値化できたことで客観的な視点から、学内の感染リスクについて議論することができるようになった。

本稿では、学内感染リスクに加えて、学内感染者数の確率も与えた。この確率には、登校できる都城高専の学生の中に感染者がいた場合に、他の学生が感染しやすくなることを仮定していない。これは、これまでの小中高校における感染事例の大半が学校内で感染者 1 人ととどまっている³⁾ことをもとにしている。もし、登校できる都城高専の学生の中に感染者がいた場合に、他の学生が感染しやすくなることを仮定すると、多くの場合は 1 人だけ感染者が出るという確率は減少し、より多くの人数の感染者が出る確率が増大する。ただし、平均値は変わらないため、本稿で定義した学内感染リスクは変わらない。

謝辞

原稿を注意深く読んで頂き、助言して頂いたことに対し、本報告の査読者に感謝申し上げます。

参考文献

- 1) 宮崎県：令和 2 年：年齢別人口構成表（令和 2 年 10 月 1 日現在）
https://www.pref.miyazaki.lg.jp/tokeichosa/kense/toke/documents/11105_20210226155627-1.xls
- 2) 鹿児島県：年齢別人口統計表一覧（令和元年）
https://www.pref.kagoshima.jp/ac09/tokei/bunya/jinko/jinkoudoutyousa/documents/56541_2020226103841-1.xlsx
- 3) 文部科学省：学校における新型コロナウイルス感染症に関する衛生管理マニュアル～「学校の新しい生活様式」～(2021.4.28 Ver.6)（令和 3 年 5 月 28 日一部修正）
https://www.mext.go.jp/content/20210514-mxt_kouhou01-000007426_1.pdf
- 4) 宮崎県：宮崎県における感染者発生状況一覧
[\[taisaku/covid-19/hassei_list.html\]\(https://www.pref.miyazaki.lg.jp/kansensho-taisaku/covid-19/hassei_list.html\)](https://www.pref.miyazaki.lg.jp/kansensho-</div><div data-bbox=)

- 5) 鹿児島県：県内・国内外の発生状況
<https://www.pref.kagoshima.jp/kenko-fukushi/covid19/hassei/index.html>

都城工業高等専門学校
研究報告第 56 号

令和 4 年 1 月印刷
令和 4 年 1 月発行

編集兼発行者：独立行政法人国立高等専門学校機構
都城工業高等専門学校

郵便番号：885-8567

所在地：宮崎県都城市吉尾町 473 番地の 1

National Institute of Technology(KOSEN), Miyakonojo College

ADDRESS:473-1 Yoshio-cho, Miyakonojo City,

Miyazaki Prefecture, Japan 885-8567

TEL (0986) 47-1109

FAX (0986) 47-1111

Research Report
of
National Institute of Technology(KOSEN), Miyakonojo College

No.56

2022

Contents

Research Papers:

- Pre-processing for Enhancing Text Data Compression Ratio
.....NAKAMURA Hirofumi, FUCHIDA Takayasu.....1
- Investigation of high efficiency in Cu_2SnS_3 thin film solar cells
...UCHIMURA Tomohiro, ARAKI Hideaki, NAKAMURA Shigeyuki and AKAKI Yoji.....13

Data and Introduction:

- An Example of Creating Collections of Compositions in Word Processor Exercise
.....NAKAMURA Hirofumi.....19
- The effect of distributed school attendance and remote learning
on reducing a risk of infection in a school.....TANAKA Mamoru.....31
