シャノンの情報理論と暗号理論の起源を振り返る

THE ORIGIN OF SHANNON'S INFORMATION THEORY AND CRYPTOLOGY

山本博資
Hirosuke Yamamoto

東京大学新領域創成科学研究科
School of Frontier Sciences, The University of Tokyo

1 まえがき

情報理論は，シャノン (Claude Elwood Shannon) が 1948 年に発表した論文 "A Mathematical Theory of Communications"[1] に始まり，また，近代的な暗号理論の研究はシャノンが 1949 年に発表した論文 "Communication Theory of Secrecy Systems"[2] から始まったことがよく知られている．しかし，それらの論文で示された多くの概念や理論は，1945 年に秘密文書 (Confidential Document) として発表された "A Mathematical Theory of Cryptography"[4] に既に含まれていたことは，あまり知られていない．シャノンの経歴は，シャノンの論文の全集 [3] に記載されているものの，シャノンが情報理論や暗号理論の着想をどのように得て研究するようになったかの経緯はほとんど書かれていない．しかし，これらの経緯はプライス (Robert Price) が 1982 年にシャノンに対して行ったインタビュー [5] で，詳しく探られている．また，岩山の著書 [6] にも比較的詳しくまとめられている．本稿では，それらの文献を基に，情報理論・暗号理論の研究がどのように始まり，"A Mathematical Theory of Cryptography" でどのような内容が発表されたかを紹介する．

2 情報理論・暗号理論誕生以前(1945 年以前 [3]–[6])

シャノンは，1936 年にミシガン州立大学で電気工学と数学の学士号を得たあと，MIT の大学院（電気工学専攻）に進学し，そこで，微分解析機の研究をしていたブッシュ(Vannevar Bush) の下で助手 (research assistant) に着いた．その微分解析機を制御するリレイ回路の設計から，ブール代数による論理回路設計法を示した有名な修士論文 "A Symbolic Analysis of Relay and Switching Circuits" の研究を行うようになった．また，ブッシュの勧めにより，電気工学専攻から数学専攻に転向すると共に，遺伝子の機能がスイッチングのように働くことから，代数が遺伝学に有用であるというブッシュの示唆により，博士論文 "An Algibra for Theoretical Genetics" の研究を行っている．なお，修士号と博士号は，共に 1940 年に取得している．

シャノンの通信に関する研究の最初の記録は，1939 年 2 月にブッシュ宛の手紙の中に残されており，時間と帯域幅，雑音と歪みのトレードオフ関係について記されている．1940 年に 9 月にプリンストン高等研究所の特別研究員になり，そこでヴァイル（Hermann K. H. Weyl）と情報と情報測定に関する議論を行っている．

1941 年にシャノンはベル研究所に入所したが，そのころには，情報を確率過程としてモデル化する研究を開始している．1943–44 年ごろに，情報および通信の諸問題について研究を行い，その後に暗号理論の研究を開始している．シャノンの情報理論の研究は，ハートレー (Ralph V. L. Hartley) の論文（雑音のない通信路の送信情報量が通信路の帯域幅と通信路使用時間に比例することを示した）を基礎としているが，シャノンは通信路の雑音や最適な符号化法を考えている点で，大きな進歩となっている．

また，この時期に，高射指揮装置 (anti-aircraft directors) の研究チームに加わり，1945 年に，国防研究委員会の報告書として，ブラックマン (Ralph B. Blackman) とボード (Hendrik W. Bode) と共著で，"Data Smoothing and Prediction in Fire-Control Systems" を発表している．これらの内容は，通信システムにおいて信号を雑音から分離する問題と類似点を持っている．また，1943 年に暗号システムの評価に協力するために渡米していたチューリング (Alan Turing) と出会い，チューリングマシンや人間の脳の働き，マシンでできることなどについて議論を行っている．これは，後にコンピュータチェスやテセウス (Theseus，迷路を解くネズミ）などのシャノンの人工知能研究に繋がる原点と思われる．

3 A Mathematical Theory of Cryptography

1945 年，シャノンは情報理論と暗号理論に関する初めての論文 "A Mathematical Theory of Cryptography"[4] を発表している．この論文は当時機密文書 (Confidential) であり，一般には公開されなかったが，現在は，国際暗号学会 (IACR, International Association for Cryptologic Research) のホームページ (https://www.iacr.org/museum/shannon45.html) から自由にダウンロードすることができる．1949 年に発表された "Communication Theory of Secrecy Systems"[2] の第 1 ページ脚注に，[2] の内容は，元々[4] で発表されたものであり，[2] の出版時点で機密でなくなっていることが書かれている．しかし，[4] の内容は，[2] の内容だけでなく，"A Mathematical Theory of Communications"[1] の内容も多く含んでいる．以下では，文献 [4] の内容を概観する．

文献 [4] は，全部で 132 ページ（文章 116 ページ，図 16 ページ）からなり，次のような構成となっている．

MEMORANDUM FOR FILE

Introduction and Summary

PART I Foundations and Algebraic Structure of Secrecy Systems

1. Choice, Information and Uncertainty
2. Language as a Stochastic Process
3. The Series of Approximations to English
4. Graphical Representation of a Markov Process
5. Pure and Mixed Languages
6. Information Rate and Redundancy of a Language
7. Redundancy Characteristic of a Language
8. Secrecy Systems
9. Representation of Systems
10. Notation
11. Some Examples of Secrecy Systems
12. Valuations of Secrecy Systems
13. Equivalence Classes in the Key Space
14. The Algebra of Secrecy Systems
15. Pure and Mixed Ciphers
16. Involutory Systems
17. Similar and Weakly Similar Systems

PART II Theoretical Secrecy

Introduction

18. Perfect Secrecy
19. Equivocation
20. Properties of Equivocation
21. Key Appearance Characteristic
22. Equivocation for Simple Substitution on an Independent Letter Language
23. The Equivocation Characteristic for a "Random" Closed Cipher.
24. Application to Standard Ciphers
25. Solving Systems Using Only N-Gram Structure
26. Validity of a Cryptogram Solution
27. Ideal Secrecy Systems
28. Examples of Ideal Secrecy Systems
29. Multiple Substitute Ideal Systems
30. Equivocation Rate
31. Further Remarks on Equivocation and Redundancy
32. Distribution of Equivocation

PART III Practical Secrecy

33. The Work Characteristic
34. Generalities on the Solution of Cryptograms
35. Statistical Methods
36. The Probable Word Method
37. Mixing Transformations
38. Ciphers of the Type $T_k H S_j$
39. The Compound Vigen[e]re
40. Incompatibility of the Criteria for Good Systems

Appendix 1 Deduction of $-\sum p_i \log p_i$

Appendix 2 Proof of Theorem 4

Appendix 3 Equivocation of Message for Random Cipher

Appendix 4 Key Appearance in Simple Substitution with Independent Letters

Appendix 5 A Theoretical Case Where All Invariant Statistics of E Are Independent of K.

Appendix 6 Maximum Repetition Rate in Compound Systems for a Given Total Key

1〜7 節では，エントロピー，同時エントロピー，条件付きエントロピー，マルコフ過程のエントロピーレート，英語の近似や冗長度など，[1] で取り扱われた内容が既に含まれている．また，一箇所 (72 ページ) だけであるが "Information Theory" という言葉が使われているとや，単位として bits の代わりに「alternatives」という用語が使われていることなども興味深い．[1]

8 節以降は，暗号理論を取り扱っているが，PART I〜III と各節のタイトルや内容は，論文 [2] の内容にほぼ一致している．つまり，[2] で示された重要な概念である完全秘匿 (perfect secrecy)，判別距離 (unicity distance)，理想秘匿システム (ideal secrecy system) などの概念を含む暗号理論が，1945 年の時点までに，エントロピーや条件付きエントロピーなどの情報理論のツールを用いて，構築されていたことがわかる．

参考文献

[1] C. E. Shannon, "A Mathematical Theory of Communications," Bell System Technical Jounal, vol. 27, no.3, pp. 379-423, no.4, pp.623-666, 1948.

[2] C. E. Shannon, "Communication Theory of Secrecy Systems," Bell System Technical Jounal, vol. 28, no. 4, pp.656-715, 1949

[3] Edited by N. J. A. Sloane and A. D. Wyner, "Claude Elwood Shannon Collected Papers," Wiley-Interscience, 1993

[4] C. E. Shannon, "A Mathematical Theory of Cryptography," A confidential Report, Sept. 1, 1945

[5] C. E.Shannon, "An Oral History Conducted in 1982 by Robert Price," IEEE Hirostory Center, Hoboken, NJ, U.S.A. (http://ethw.org/Oral-History:Claude_E._Shannon)

[6] 岩山知三郎, "インターネット・サイエンスの歴史人物館," 第 2 章「クロード・シャノン」, 株式会社インプレス R&D, 2012 (ISBN：978-4-8443-9528-7)

[1] bit という用語も一箇所であるが使われている．