暗号・情報保全史特論
History of Cryptograph and Signal Security Advanced Course

第5・6回: 多表式暗号

佐藤永欣

多表式暗号:
単一換字暗号の換字表が複数ある方式

- 換字表は一定の規則で切り替える
  - 初期には順番通り切り替え→後にキーワードで切り替え→ランダム切り替えに変化
- あらかじめ用意した乱数に従って切り替え(乱数多変式)
  - ヴィジュネル暗号、M-209(アメリカ)、バーナム暗号
  - 換字表の秘匿性はあまり重要ではない。暗号書・暗号機などが盗まれる・遺棄される可能性がある陸戦部隊の前線向けによくつかわれた
  - 乱数表が漏洩するとほぼ役に立たなくなるが、乱数表は頻繁に更新される
  - 乱数を大量に消費する。乱数の周期が十分に長くなければ安全ではない
    - ヴィジュネル暗号のキーの周期=1が極端な例。シーザー暗号と変わらない
- 順次切り替え(順次多変式)
  - エニグマ、パープル; 退化のように思えるが機械式暗号機の実装上やむを得なかった
  - 換字表の秘匿が重要。暗号機が盗まれたりしない環境むけ。上級司令部、艦艇、大使館
  - 多表の実現方法・多表の切り替え方法により暗号文に特徴が現れ解読の糸口になることがある
- 換字表の切り替えに周期性がないのはワンタイムパッドだけ
  - 周期性がなければ解読不能になる(シャノンによる証明)が、平文と同じ長さの乱数が必要
    - 疑似乱数も×

多表式暗号:
単一換字暗号の換字表が複数ある方式

- が、多表式暗号が広く普及するのは20世紀に入ってから
- 主な原因
  - 暗号化・復号化の作業が複雑すぎる
  - 比較的ローテクな道具で作業を楽に実施できるM-94暗号の場合でもかなりつらい
    - M-94暗号については概要を紹介
  - 暗号文が音声言語と無関係な綴り字になる→書きにくい・読みにくい
    - 漢字の部品(偏・冠・旁など)がでたらめに入れ替わって組み合わされるのを想像すればいい???
- 単語単位で別の単語に変換する方式が20世紀になるまで暗号強度を保つための主要な方法として使われ続けた

コードブック+ヴィジュネル暗号の時代

- 外交暗号などの実用のレベルでは、コードブックによる暗号化+ヴィジュネル暗号の時代が20世紀に入ったころまで続いた
  - 19世紀半ばにはコードブックだけを使った暗号は安全ではなくなった
  - ナポレオンやルイ14世の暗号を解くデモンストレーション
- 手作業で暗号化・復号化をするのであまり複雑なことはできない
  - 複雑にしすぎると暗号化時のミスや電信で送るときのミスなどが原因で復号化できなくなる
  - コードブックに秘匿性を依存する時代が長く続いたのも複雑なことを嫌ったから?
  - タイプライターの発明が19世紀半ば→暗号化・復号化を機械化できる萌芽
- (小手先の技に近い)解読対策がなされた
  - ヴィジュネル暗号の鍵の反復長を推定する方法が発表される(1863: カシスキー)

コードブックの延命

- コードブックによる暗号化結果にランダムな文字・数字を加算する方法が20世紀に入って考え出された
  - Ex) コードが5桁数字→5桁の乱数を加算。6桁目は無視
    - コードがアルファベットの場合でも同様。A+Bを1+2とみなしてA+B=Cとする。ただしmod 26の世界での計算になる。シーザー暗号の計算とおなじ
  - 実質的にはコードブックを水増ししているのと同じ
    - アルファベット同士で足し算をする発想は19世紀末～20世紀初頭までなかったらしい。今は当たり前なのに
  - 加算する乱数は送信者、送信日などによって決める
  - 後には乱数表から順番に1文字ずつ取り出し、加算する乱数を順次変更することも行われた
    - ここまでくると機械式暗号と大差ない
    - 手作業では暗号化と復号化が非常に大変
    - ヴィジュネル暗号の類に乱数を加算することも行われた

古典暗号の解読対策

- 分かち書きなど
  - 平文の単語の区切りの空白を除去
  - 一定の文字数（5文字が多い?）ごとに区切り直し。空白除去だけで区切り直さないこともある
- 分割転置: 通信文の中での文章のシャッフル
  - 通信相手や通信内容は比較的決まっている
  - 通信文の先頭にこれらが来ることが多い → 解読のヒント(クリブ)
  - 1通の通信文を数ブロックに分割 → ランダムな順序で送信。受信側で意味が通るように組み立てなおす
- 頻出語句のコードによる置き換え
  - コード表を定期的に更新して頻出語句が解読のヒントにならないようにする
- パラフレーズ
  - 相手国の外務省からの通知文、新聞記事など、平文が公知 or 既知の場合、意味が変わらない程度に別の文章に書きかえる

機械式暗号の時代

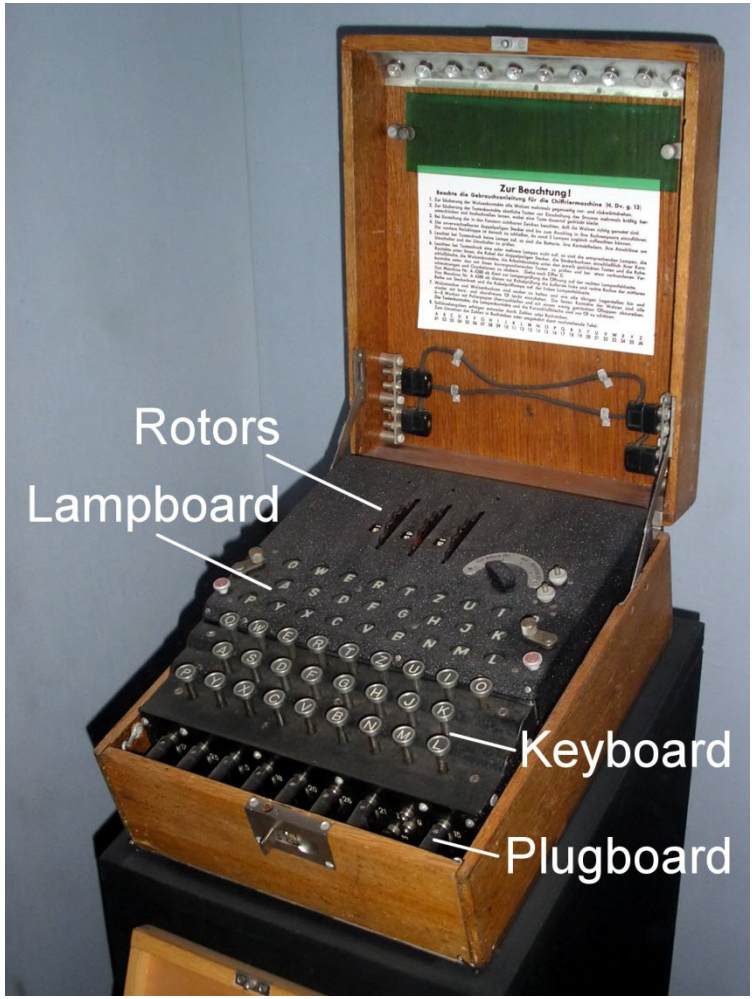
- 手作業での表引きから道具・機械へ
  - アルベルティの暗号円盤
  - ホイットストーンの円盤
  - サンシール暗号計算尺
  - M-209暗号円盤・ストリップサイファ
  - クリハ暗号機
- 機械・器具は使うようになったが暗号書と併用
  - 日本海軍の場合: コードブックでコード化→換字表で1対1変換→97式暗号機かD暗号書でサイファ化→電報で送信
  - 暗号書も暗号機も重要気密なのでかかわる人を減らしたい→同じ海軍組織内でも機密度や方面・担当業務別に暗号システムがあった

機械式暗号の時代

- タイプライタの発明
  - ミシンを作れる工業レベルになると機械式タイプライタを作れる
  - レミントンタイプライター: 現在も使われているキーボード式(1860年ごろ)
- 暗号化と復号化を自動的に行う
  - タイプライタのキーボードと、暗号化後の文字を印字 or ランプで表示するタイプ: 第一次大戦後に実用化
    - Enigma
    - 97式印字機(各種)
    - M-209
  - 携帯用にスライド式の変換表を使うタイプ
    - M-94 strip cipher
    - M-138A cylinder cipher

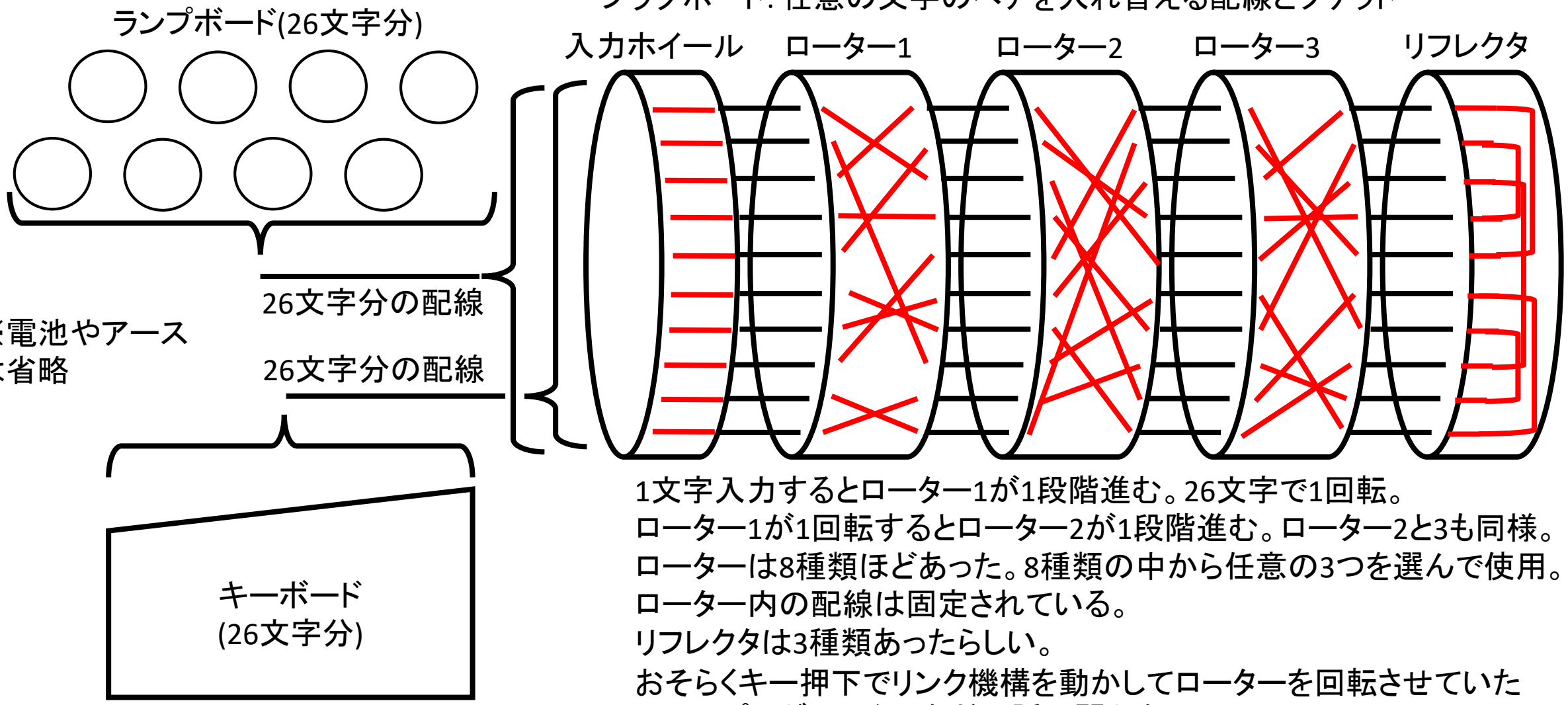
Enigma

- 第2次世界大戦中にドイツが使用していた暗号機
  - 原型は1918年開発→1925年に軍が採用
  - ドイツ政府、ドイツ帝国鉄道も採用(軍用とは配線や暗号強度が異なる)
  - 商用として民間向けバージョンもある
  - スイス軍もなぜか使用している
- 多表式暗号を回転するローターによる回路のつなぎ替えで実現
  - リフレクタで信号を折り返す。3個のローターを計6回通過→同じ文字に暗号化されないという欠点
  - その代わり暗号化と復号化の切り替えが不要
  - ローター式なので斜行特性がある(隣接する文字の変換に互いに関係がある。平文で連続する文字が来るとローター1のハミング距離がわかる)



Enigma

キーボード→入力ホイール→ローター1～3→リフレクタ→ローター3～1→入力ホイール→ランプボードと電流が流れる後にキーボード→入力ホイールと入力ホイール→ランプボードの部分にプラグボードが追加された。
プラグボード: 任意の文字のペアを入れ替える配線とソケット



1文字入力するとローター1が1段階進む。26文字で1回転。
ローター1が1回転するとローター2が1段階進む。ローター2と3も同様。
ローターは8種類ほどあった。8種類の中から任意の3つを選んで使用。
ローター内の配線は固定されている。
リフレクタは3種類あったらしい。
おそらくキー押下でリンク機構を動かしてローターを回転させていた(ステッピングモーターなどの話は聞かない)。

Enigma

- 使い方
  - 暗号書を参照して日鍵・時鍵にもとづいて
    - プラグボードを設定
    - ローター1～3を8種類の中から選択
    - ローターの繰り上り位置を決める(4種類から選べる? ほとんど変更されなかったらしい)
    - リフレクタを選択
  - ローターの開始位置を決める
  - 1番目の平文の文字のキーを押す→暗号文の1番目の文字のランプが光るのでメモ(以下繰り返し)
  - メモした暗号文を電信で送る
    - 電文が何通もあるときは同じ開始位置・設定変更なしで送るなどの手抜きをした模様
    - ローターの開始位置: 電文で送った説と規約で決まっていた説がある
      - 時期によるのか不明

Enigma

- 斜行特性: ローター1個だけで考えてみる

	a	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**	**j**	**k**	**l**	**m**	**n**	**o**	**p**	**q**	**r**	**s**	**t**	**u**	**v**	**w**	**x**	**y**	**z**
0	E	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P	A	I	B	R	C	J
1	J	E	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P	A	I	B	R	C
2	C	J	E	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P	A	I	B	R
3	R	C	J	E	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P	A	I	B
4	B	R	C	J	E	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P	A	I
5	I	B	R	C	J	E	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P	A
6	A	I	B	R	C	J	E	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P
…	…																									
25	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P	A	I	B	R	C	J	E
26	E	K	M	F	L	G	D	Q	V	Z	N	T	O	W	Y	H	X	U	S	P	A	I	B	R	C	J

Enigma

- ローター1個だけで考えると色々まずいことが起きる
  - ローターがn段進んだ時、平文のハミング距離がn→出力が同じ
  - n文字でローターがn段進む
  - n文字離れたところに同じ暗号文字がある→平文のハミング距離はn
  - 26段階進むと次のローターが1段階進む→n<26ならば常に成立
  - 26≦n<52, 52≦n<78,…で同じことが起きる
- 隣接文字のハミング距離のクリブ(のようなもの)を使う
  - ローターの配線を推定できる
  - 実際にはローター3個+往復分なので複雑だが原理的には同じ
  - 26文字まではほかのローターは変わらない
  - 26文字よりも早くほかのローターを回すことにしても原理的には大差ない
    - 発見しにくくはなるだろうがローター3個分の周期は最長でも26^3=17576
    - 例えば5文字ごとに次のローターを回す

Enigma

- 運用上の問題
  - 定型の通信文を使いすぎた
    - 気象観測データ
    - 書き出しがワンパターン: Heil Hitler、Heil Führerで毎回始めるetc
    - 長いクリブがある: ドイツ語の特性上仕方ない? (名詞と名詞を連結して新しい名詞を作る)
  - 長文の通信文を同じ暗号鍵で送った
    - 1万数千文字の命令書など: 大量のクリブを与えるのでパラフレーズしたうえで分割すべき
  - ローター開始位置を恣意的に決めた
    - ランダムにしたつもりでもランダムにはならない(人間だもの
    - 面倒だといってAAA、ABCのような開始位置を多用する通信員がいた
  - 後からプラグボードを追加した
    - かえって暗号強度が下がる(最初からあれば問題ない)
  - 配線が異なるローターの作成・配布を怠った

Enigma

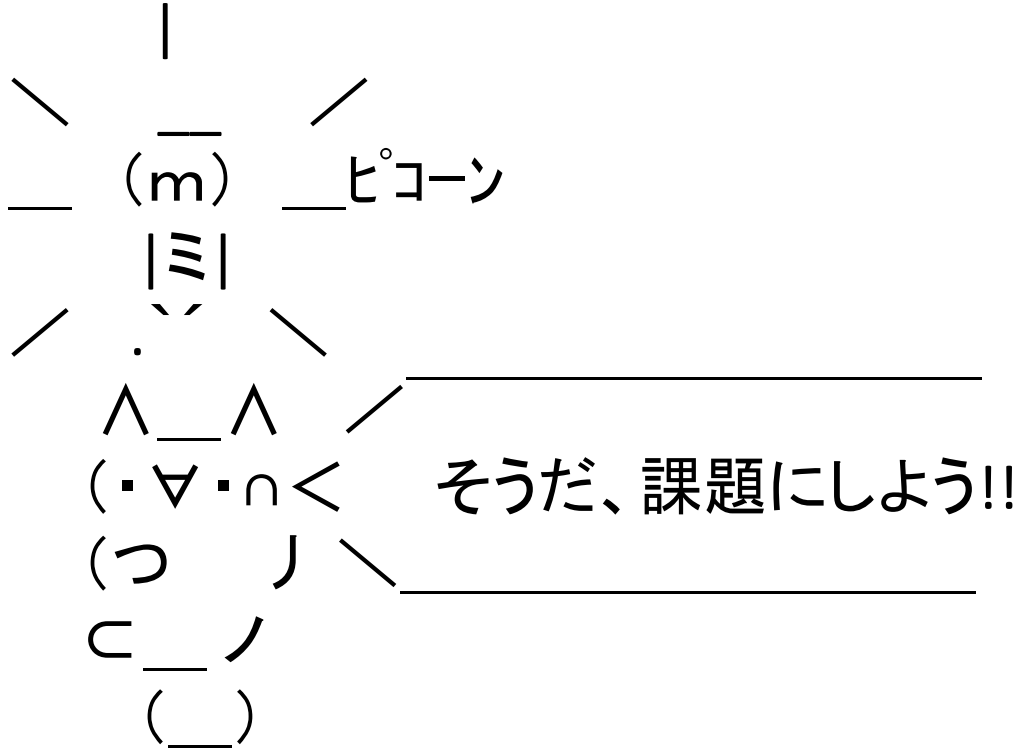
- イギリスによる解読方法
  - ポーランドが開始していたEnigmaの解読を継承(ポーランド自体は東西からドイツとソ連に攻め込まれて消失)
- Bombe: 次の組み合わせを全数探索。クリブを頼りに判断
  - ローターの選択
  - 各ローターの繰り上り位置の設定
  - プラグボードの設定
  - ローター開始位置
- 1台のbombeに36台のエニグマ模造機相当品が内蔵
  - 平文のクリブを見つけるとbombeが止まるようになっていた

Enigma

- 模造機(というかプログラム)について
  - いっぱいネットにおちてるから探して(投げやり
    - ネットに落ちているものはキー入力とランプのUIのものばかりなのでイギリス情報部ごっこをするには使いにくい
- 配布するEnigmaシミュレータ
  - Rubyで記述
    - FreeBSD環境でしかテストしていないがLinuxでも動くはず
    - テキストの読み書きと標準出入力しか使っていないのでWindowsでもターミナルから使えるはず
  - 歴史上存在したローターとリフレクタの設定から選択できる
    - 一応プラグボードの処理はしているがA→A, B→B, …, Z→Zの変換にしてある
  - 標準入力から平文、標準出力から暗号文(またはその逆)
    - Enigmaには暗号化・復号化の切り替えはない

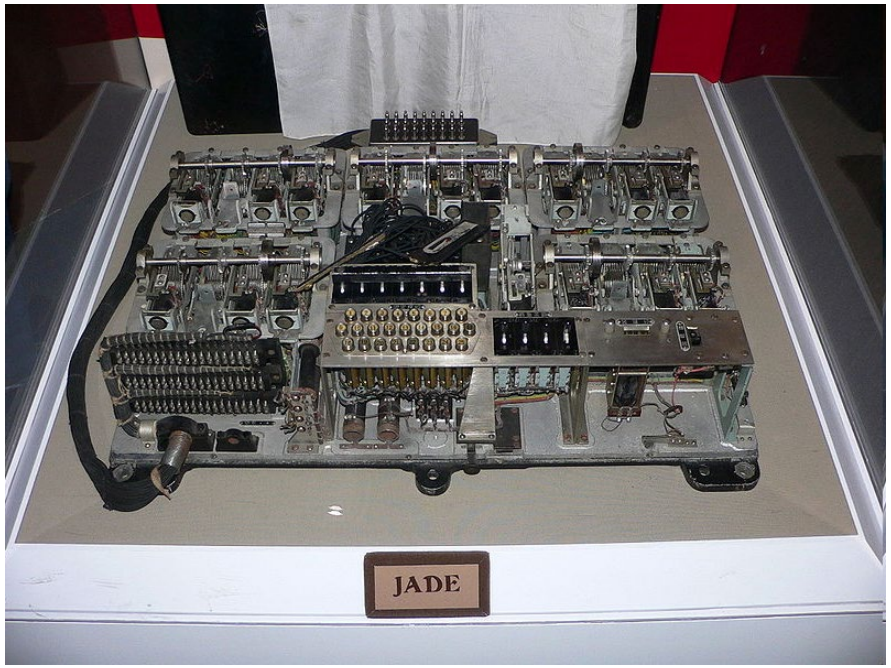
Enigma

- Bomeのシミュレータも作りたいが…



九七式暗号機(海軍用・外務省用)

- 類似する日本の暗号機にアメリカによるコードネームがついている
  - 虹の色の名前
- 九七式印字機一型/二型(海軍艦艇用)
  - JADE
  - カタカナの暗号化と復号化
  - 暗号書(コードブック)と組み合わせて使用されていた
  - 1941年の真珠湾攻撃直前から1944年ごろまでに使用された
  - 1942年中ごろには一部が解読されていた
  - かなりの重量があるが、艦船に積載したままなのであまり問題にならない



JADE(米軍による捕獲品)

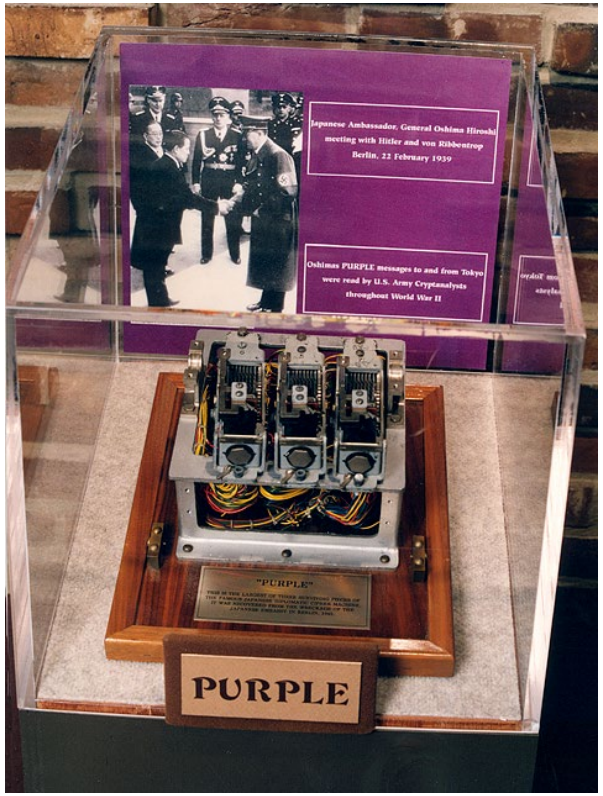
九七式暗号機(海軍･外務省用)

- 外務省用の暗号機は海軍の暗号機をもとに開発された
  - 電報用に欧文のみ
  - AIUEOYとその他の文字で暗号化機構が異なる
- 運用が原因で1941年初めには解読されていた
  - RED暗号とPURPLE暗号で同じ平文を送信
  - 平文の冒頭がワンパターン
  - 句読点コード表が盗まれた
  - プラグボードの接続変更を怠った

海軍側は九七式暗号機の暗号強度の限界を認識していた模様

作戦部隊は九七式暗号機ではなくD暗号を常用していたらしい

九七式欧文印字機三型（外務省用）
PURPLE（アメリカによる模造機）



PURPLEの米軍にベルリンで捕獲された部品
（ロータリーラインスイッチ）

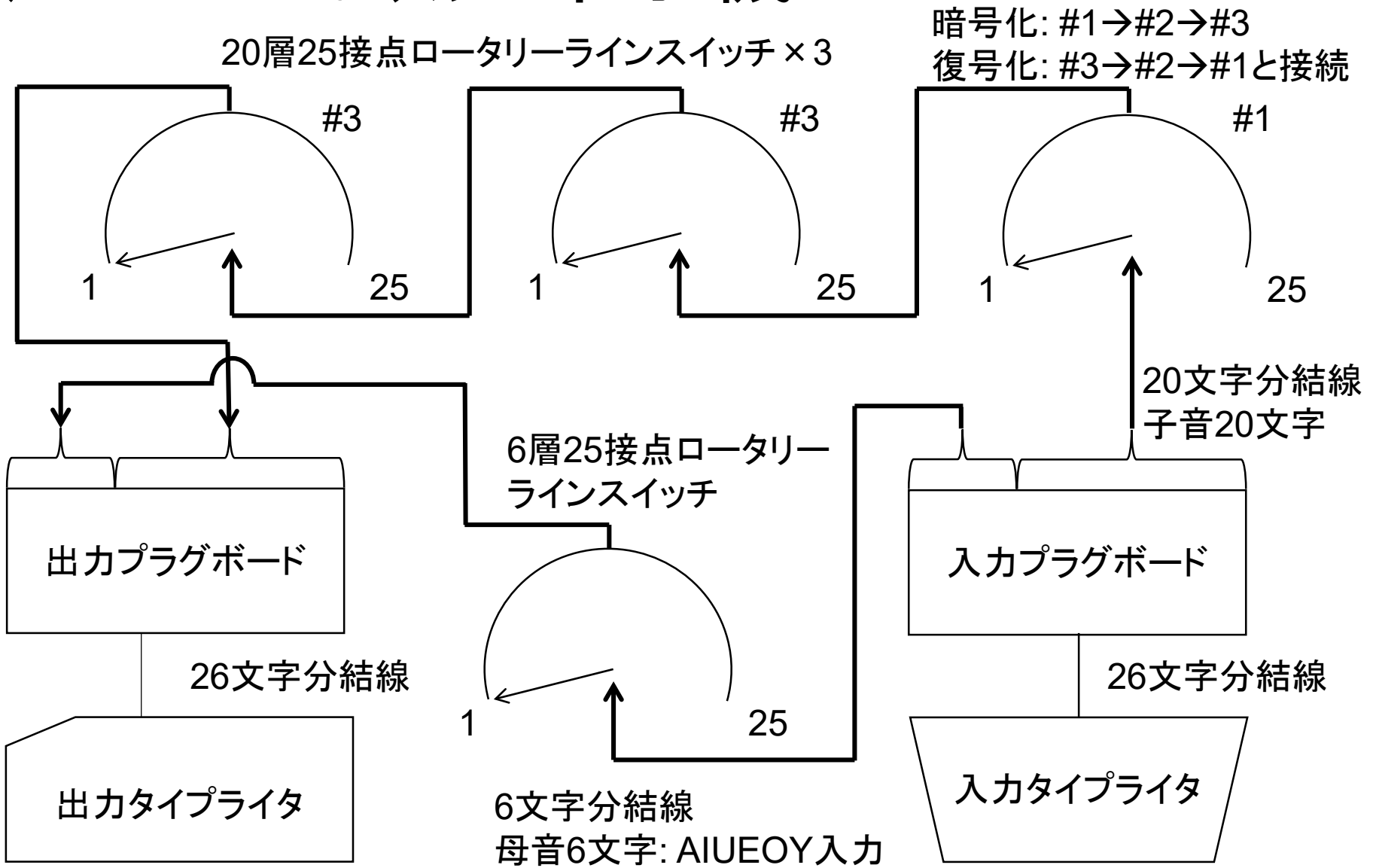
外務省用九七式欧文印字機

- アメリカ側コードネーム PURPLE
  - 日本の暗号に虹の色名のコードネームをつけていた
- 海軍用九七式欧文印字機をもとに開発
  - 開戦前の日米交渉や最後通牒をこの暗号機でやりとり
    - アメリカには解読されていた
  - 皇紀2597年に完成したので九七式
- 電気式タイプライタ2台を暗号機を介して接続
- 片方で平文を入力すると、もう片方から暗号文が出力される
  - 入力用タイプライタのキーボードと出力用タイプライタの印字ハンマー駆動機構が電気的に接続される
  - 暗号機はこの接続を切り替える
- ほかにも九七式暗号機があるので注意
  - 陸軍用: エニグマと同じローター式、海軍の九七式とは無関係
  - 海軍用: 海軍艦艇用(JADE, 97式1型と2型)、海軍駐在武官用(CORAL, 97式3形)

外務省用九七式欧文印字機

- A～Zの26文字を取り扱える
  - 句読点や記号は2~3文字のコードで表現
- 構成要素
  - 入力用タイプライタ
  - 入力用プラグボード
  - ステッピングスイッチ(ロータリーラインスイッチ)
    - 大昔の電話交換機に使用されていた、電話回線の切り替えスイッチ。パルスを入力すると入力を次の接点に切り替える
    - 母音用6層25接点スイッチ
    - 子音用20層25接点スイッチ
      - 3個直列に接続。スイッチング速度を3種類から選べた
  - 出力用プラグボード
  - 出力用タイプライタ

外務省用九七式欧文印字機



http://cryptocellar.web.cern.ch/cryptocellar/pubs/PurpleRevealed.pdf による

外務省用九七式欧文印字機

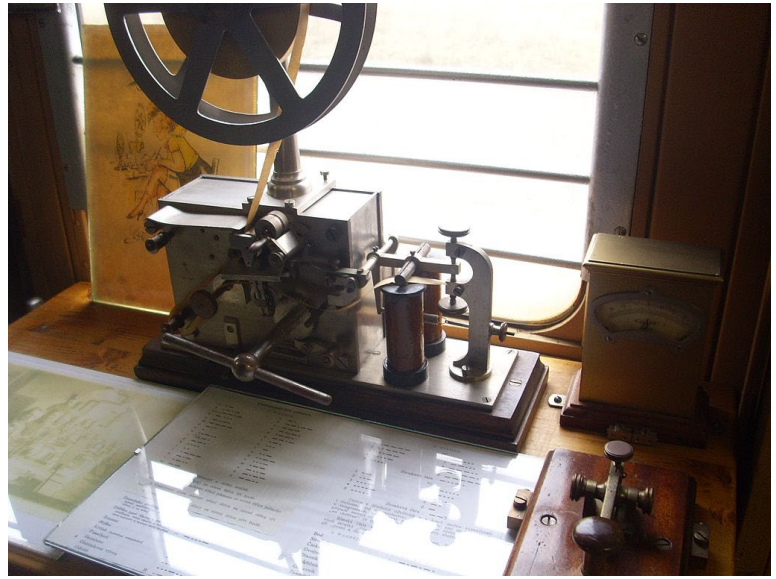
- 特長
  - Enigmaのようなローター式と違い、ステッピングスイッチの入出力間を全く無関係に接続できる。
    - N文字目とN+1文字目の換字表を無関係にできる
    - 斜行特性がない
  - 入力と出力のプラグボードを無関係に設定可能
- 弱点
  - 母音と子音で暗号化プロセスが異なる(という説が有力)
    - Artificial words対応仕様のため。必ず母音が入るので電報料金が安い
  - 母音側6文字(AIUEOY)は別の6文字に変換される
    - 母音は使用頻度が高いので頻出文字ができる→母音以外の文字に変換しても頻度分析で母音字を推定できてしまう
    - 6!=720通りの組み合わせしかない
  - 子音側スイッチ間の接続が固定。Enigmaのようにローター入替で換字表を変更できない

外務省用九七式欧文印字機

- 運用上の問題点
  - プラグボードの配線が長期間更新されなかった
  - 入力と出力のプラグボードの配線は同じか、わずかに違うだけだった
  - 当初、母音AEIOUYをAEIOUYにしか変換しなかった
  - 子音用ステッピングスイッチの配線が変更されなかった。すなわち、同じ暗号表を使い続けた
  - 対照平文がわかっている暗号文を送信した
    - 相手国から渡された文書・新聞記事etcをパラフレーズも分割転置もせずに送信
    - 「いずれ公知になる・すでに公知の事項」という意識があったらしく大変雑な運用をした
    - そもそも外務省に暗号の専門家がいなかった: 暗号課長も文系出身の外交官で数学的な話は通じなかった
  - 同じ平文をRED暗号機とPURPLE暗号機の両方で送信した。REDは解読されていた
    - 大使館にはPURPLEがあるが公使館・領事館にはREDしかない等
  - 電文の出だし・形式が毎回同じ
    - 型式や格式にこだわる役所にありがちではある。せめて分割転置していればマシだったかも?
  - 「先頭に適当な文字を入れよ」のような通信員向け指示をそのまま送信

補足: 電信

- 一組の電線の電流オンオフで通信
  - 近世からあった腕木式の通信を置き換えた
  - モールス電信機: 1836年発明
    - 他にも方式がいくつかあるが最も普及したのがモールス式
    - モールス信号を音で受信するか紙テープにペンを押し付けて受信する
- 電報
  - 電信を使った通信手段で一般的だった
  - 公衆用は電報局から配達に来る
  - アルファベット・数字(・カナ)を送れた
  - 文字数で料金が決まる



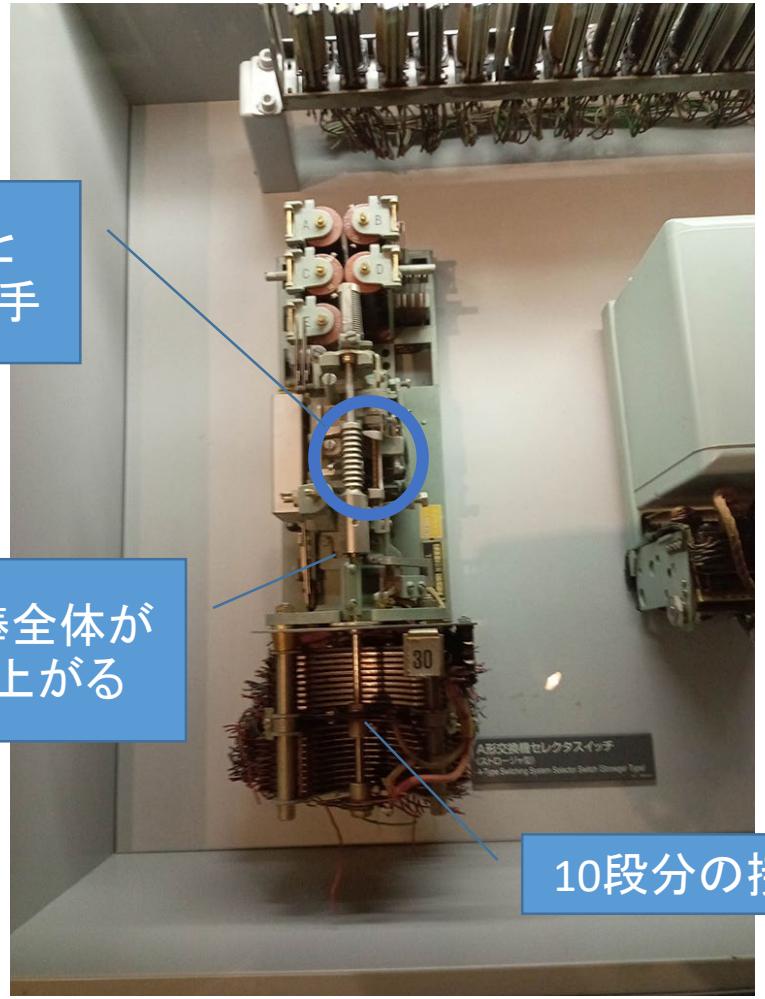
補足: ロータリーラインスイッチ

- アナログ電話の自動交換機で回線交換をするためのスイッチ
  - パルス式回線用
  - パルスが1つ入力されると電磁石で1ステップ進む
  - 電話番号1ケタに対してスイッチ一つが対応
  - 234-5678→4パルスで4を選択、次のスイッチに5パルス入力して5を選択…
- パルス回線⇔トーン(プッシュ)回線
  - 2本の電話線の短絡・切断で電話番号を伝える方式
  - 受話器を上げると電話線が接続
  - 受話器を置くと切断、パルス送出でも瞬間的に切断(交換機に48Vかかるので電磁石を駆動できる)

1. 受話器を上げるとツーっという音が聞こえる
2. 数字の内側の穴に指を入れて右にまわす
3. 指を離すとダイヤルが一定速度で元に戻る
4. ダイヤルが戻るとき、数字に対応した数のパルスが交換機に送出される

補足: ロータリーラインスイッチ

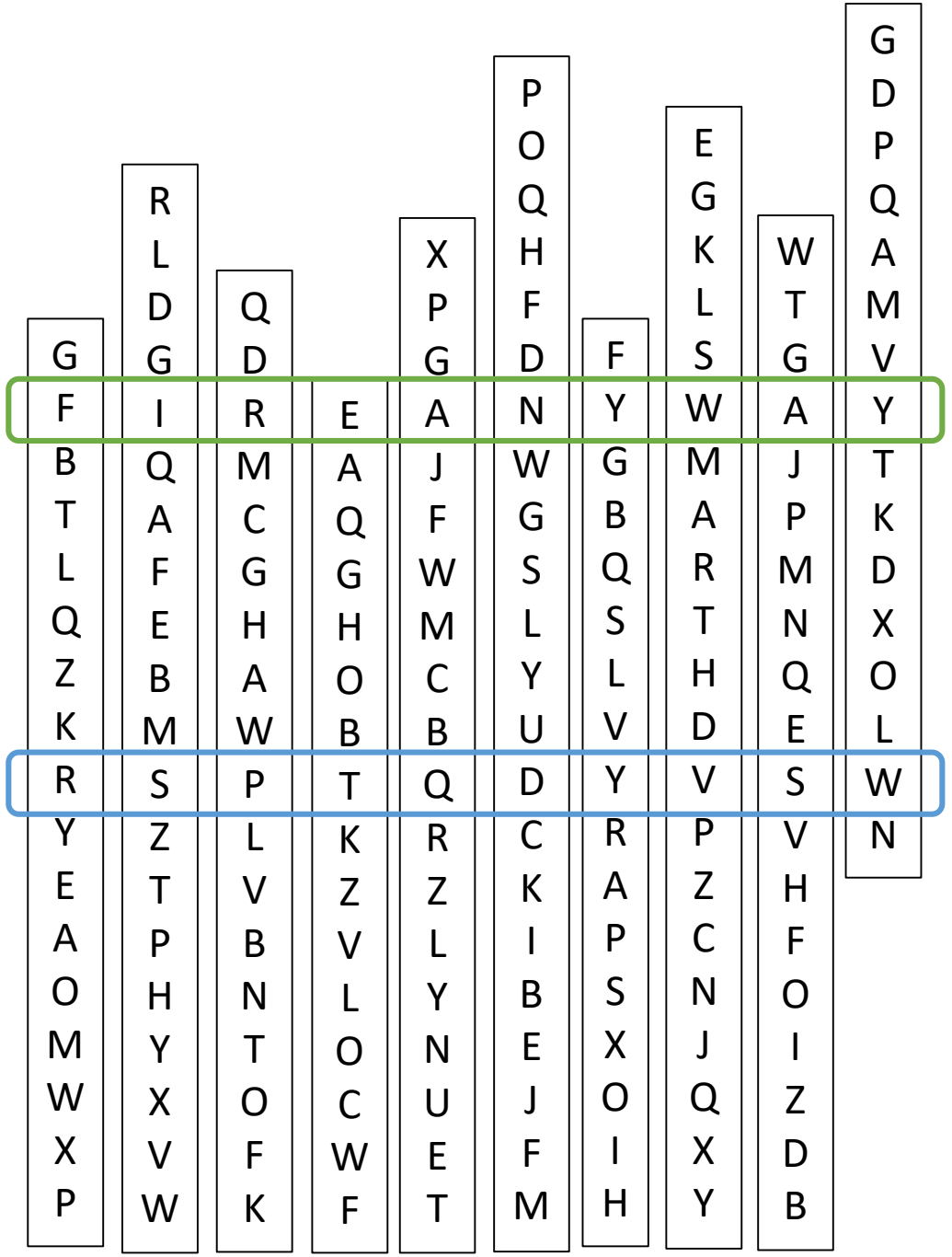
- 戦後のロータリーラインスイッチ
  - 電話線には48Vが供給されている
- 受話器を上げる
  - フックスイッチが短絡→マイクとスピーカーが交換機に接続→2本の電話線の電圧が6Vぐらいになる
  - ダイヤルを回す→ダイヤルが戻るときに電話線の接続が一時切れる→48Vのパルスが出る
- 交換機側
  - パルスの入力1回でラチェットのロックが1回外れる→接点が1段階上がる



A型交換機用ストロージャスイッチ
NTT技術史資料館(武蔵野研究所)

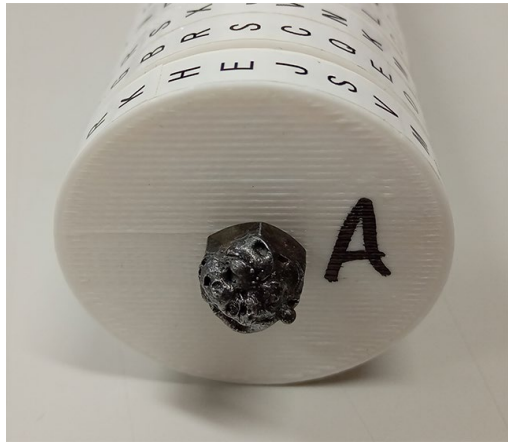
Strip cipher (M-94)

- スライド式の多表
  - ストリップには縦にアルファベット(+数字)が並んでいる
  - ストリップを決められた順番に横に並べる
    - 26本セットなどになっている
    - 文字配列とは別にID番号がついている
    - 順番は日時・部隊名などで決まる
  - ストリップを平文の文字に合わせて縦にスライドさせる
  - N文字下を読み取って暗号文とする
    - Nは電文で送ったり送らなかったり
- 回転できる円筒の外周に文字を刻んだものもある。原理的には同じ
  - M-138A cylinder cipher

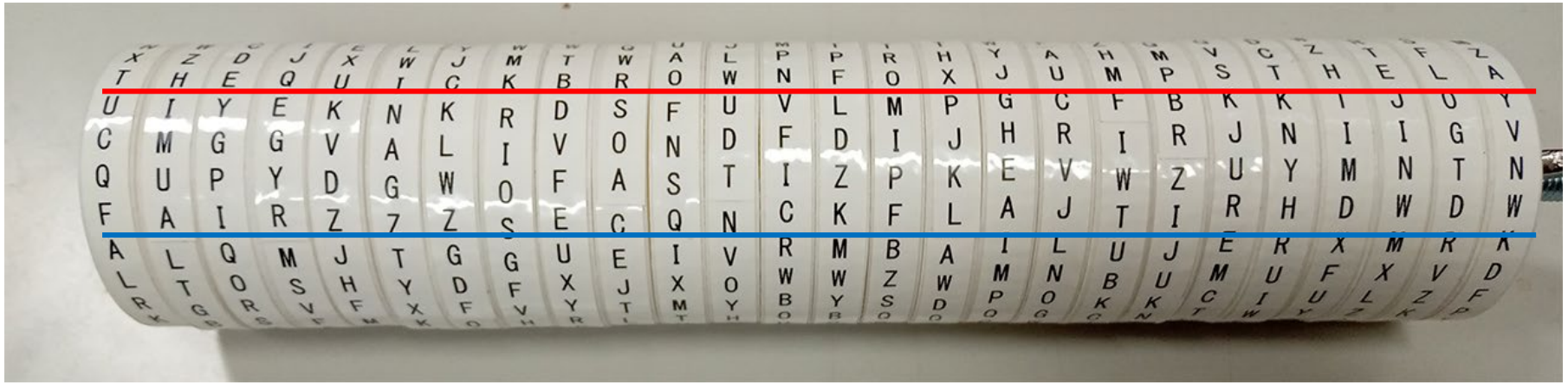


Strip cipher (M-94)

- 円筒形の模造品を作ってみた
  - 棒ネジにとおしてナットで固定
  1. ディスクA～Zを暗号書に従って並べ替える
  2. ディスクを回して平文に位置を合わせる
     the quick brown fox jumps the la(zy dog)
  3. 暗号文を読み取る: オフセット4のとき
     FAI RZZZS ECQNC KFL AJTIR HDW DW
  4. 26文字を超える長い暗号文は2に戻ってディスクを回して3を繰り返す



こんな感じで各ディスクに名前がついている



次回予告: 多表式暗号の解読技法

- カシスキーの方法によるカギの反復の長さの推定
- ケルクホフによる換字の状態を合わせる方法
- フリードマンの一致反復率

課題用プログラムなど

- ヴィジュネル暗号
  - vigenere.rb 第一引数にキーワード。標準入力から平文を入力

- Enigma
  - 歴史上存在したローターとリフレクターの配線を再現してある
  - プラグボードも一応実装
  - 詳細はengmalib.rbのコメントを参照
  - 使いかた: enigma.rb ローター#1 ローター#2 ローター#3 ローター開始位置
    - 標準入力から暗号文・平文を入力、標準出力から平文・暗号文を出力する
    - ローター開始位置は省略できる。省略しない場合は3文字のアルファベットで指定