Unit 185
Attacking Purple with cribs

If we know the settings of the Purple machine at any point in a text, we can run the machine forwards or backwards and find its settings at any other point. So for the purposes of illustrating the method of attacking it with a crib, we may as well look at cases in which we know the plaintext from the beginning of the text; in other words, a known-plaintext attack.

Let's just work an example and explain along the way. Take this ciphertext and plaintext pair from *Riders of the Purple Sage* by Zane Grey:

```
YBUMVFPNMDCXUYRRQNGEQYAQLEWKCLHDZRDQUZSJNYBBWGINJSZEYZODOAVXEOQG
SWFGFDZUQTYHOSBXTOUQXLQBUPCHOSDLEQLYZEJHZBOQQSOKOJGBCXGBNCXOHWJF
KATSEPTUUHOYIDJZDMWZBBNJWKZYSBUXXNIQAIMNZXIWZOQOGJLIUBVEZHVTZDNL
OBETYGONLXUZGXWQFPGQATHFETKNZVCSTAQWNFVZJCRQFIXYMUKPEMTMBIBEXLSR
OTNVZHIVMSYGDLINFKWZHWRECUQBSWREKFCZJSKMWURPJSDXODELJIFMAFNHWLYO
GSCBDZUDSRNSOIOMFQFSDOIXDSVEKGLOTCEWKCNPPYVVMLQJDLNFSJYFBOLFEFDK
RNNWRVXYAPXRAUPVFAMMJSZOHEYMMWWNJZGIRRZHSRWSGEELICRWXNPAHPJKRDUT
PNLCMFFWNYJAOGSJVTCGGOUHNOQMIOOXTNQUTFODYSFJLVZMTRROPPKRTYYJCTZO
UMDIOATKSFSEYY

ITWASTHEMOMENTWHENTHELASTRUDDYRAYSOFTHESUNSETBRIGHTENEDMOMENTARI
LYBEFOREYIELDINGTOTWILIGHTANDFORVENTERSTHEOUTLOOKBEFOREHIMWASINS
OMESENSESIMILARTOAFEELINGOFHISFUTUREANDWITHSEARCHINGEYESHESTUDIE
DTHEBEAUTIFULPURPLEBARRENWASTEOFSAGEHEREWASTHEUNKNOWNANDTHEPERIL
OUSTHEWHOLESCENEIMPRESSEDVENTERSASAWILDAUSTEREANDMIGHTYMANIFESTA
TIONOFNATUREANDASITSOMEHOWREMINDEDHIMOFHISPROSPECTINLIFESOITSUDD
ENLYRESEMBLEDTHEWOMANNEARHIMONLYINHERTHEREWEREGREATERBEAUTYANDPE
RILAMYSTERYMOREUNSOLVABLEANDSOMETHINGNAMELESSTHATNUMBEDHISHEARTA
NDDIMMEDHISEYE
```

The first thing we want to do is identify the sixes and the twenties. For that purpose, we tabulate the ciphertext letters of each plaintext letter.

plaintext letter	ciphertext letters
A	MADOCK
B	GFJYQPNU
C	OD
D	KCOMDA
E	NXQSBEZVGUYTHLFWIPJR
F	QFSBWZUHNY
G	JXBWIQLET
H	PREZSUBYIGNFVJX

I	YNGTSXQWHZJFEPLI
J	
K	OM
L	YSHLIBGPRNXW
M	MCDAOK
N	UNYXBHLJPIETFSWZRVQ
O	DOKCMA
P	XFEWVQU
Q	
R	EHIQZLXJTVRNSGPY
S	VQRJBHFSTUWENGLYXI
T	BFYGLUWZETPHQXVSRI
U	WNQXZTRFHJ
V	EUG
W	URQXNTJPIZSFW
X	
Y	LZWQBFNJY
Z	

Any letter that has more than six images (what mathematicians call the thing you get after applying a mapping, such as the encryption function) is one of the twenties. Furthermore, any letter with an image that is one of the twenties is also one of the twenties. And any letter that is an image of a twenty is also a twenty. The sixes must form their own closed group. We now know the sixes and the twenties, but not in any particular order, and definitely not in the order as mapped by the plugboard.

sixes:	ACDKMO
twenties:	BEFGHIJLNPQRSTUVWXYZ

Next we are going to determine the initial setting of the sixes switch *S*. Start by tabulating the mappings of the sixes and arrange them according to the position in the text, modulo 25 (the number of switch settings). In this table, we take the first position in the text as zero. Right away we notice that for position mod 25 equal to 2, there are two fixed points: M→M and O→O. By studying Table 184.1, we know that this means that at position 2 in the text, the sixes switch can be in position 1, 5, 8, 10, 14, 15, 18, 19, or 22. Position mod 25 equal to 14 also has two fixed points. So positions 2 and 14 (12 apart) can have sixes-switch positions 10 and 22, or 14 and 1, or 18 and 5. Position 16 has a chain of five: A→D→M→O→C. This is either a five-cycle with a fixed point (A→D→M→O→C→A and K→K) or a six-cycle (A→D→M→O→C→K→A). Sixes switch position 24 has a five-cycle plus fixed point, so position 2 in the text is consistent with switch position 10. Switch position 7 has a six-cycle, so position 2 in the text is consistent with switch position 18. Switch position 3 does not have a five- or six-cycle, so we can eliminate switch position 14 for text position 2. Suppose text position 16 is switch position 24. Then its fixed point is K, which must be 5, since the permutation is (3 4 6 1 5 2). Text position 13 has A→K, so A must be 4, because the permutation of switch position 21 is (3 4 6 5 2 1). In text position 14, switch position 22, 4 is a fixed point, so that is consistent. But now suppose that text position 16 is switch position 7. Text position 17 has a cycle of length at least three, but switch position 8 does not. This is not consistent. We conclude that text position 16 has switch position 24. This means that at the start of the text, the sixes switch is in position 8.

position mod 25	mappings of sixes	position mod 25	mappings of sixes
0	C→O A→A D→M	13	A→K D→M→O
1	D→O→A→C	14	A→A D→D O→M
2	A→D→K O→O M→M	15	A→C K→M O→D
3	A→M→O→K D→C	16	A→D→M→O→C
4	D→O M→A	17	D→O→K M→A
5	M→D→C	18	A→K C→D
6	A→D O→O M→K	19	A→O→D→K
7	M→A→O→C D→D	20	A→M D→D
8	M→M A→C D→K	21	M→C D→A
9	O→D→O M→A	22	A→A O→C D→M
10	A→M→C O→O	23	A→O M→K
11	A→O→K M→M	24	M→O→D
12	K→O→M→D A→A		

Now let's see if we can deduce the plugboard mapping for the sixes. We already know that text position 16 has switch position 24, which has a five-cycle and a fixed point, which is 5. From the table above, that fixed point must be K. So we have 5=K. We also saw that in text position 13 with switch position 21, A→K, and the permutation there is (3 4 6 5 2 1), so 4=A. At text position 2 with switch position 10, A→D→K, and the permutation is (5 4 3 1 2 6), so 1=D. Also at that point, O and M are fixed points, which are 3 and 6 (but we don't know which is which yet), so the remaining letter C must be 2. At text position 3 with switch position 11, A→M, and the permutation is (2 1 6 3 4 5), so 3=M, leaving 6=O. We can now conclude that the sixes portion of the plugboard is DCMAKO.

Now that we know the starting position of the sixes switch, we not only know its position at every point in the text, but we also know when the fast twenties switch advances (every point) and when the medium switch advances (when the sixes switch advances from 24 to 25). There are $3! \cdot 25^3 = 93{,}750$ possibilities for the initial settings of the twenties switches. This is a manageable number with modern computers. Our strategy is to try them all and find ones that are consistent with the plaintext-ciphertext pair that we have. A good place to start is with fixed points. The chance that a random permutation of twenty items has a fixed point is about 65%, which is rather high. But keep calm and carry on. The fixed points of letters in the twenties in our texts occur at these positions:

position in text	fixed letter	position in text	fixed letter
17	N	385	N
51	E	388	R
68	F	420	R
80	T	426	W
85	L	429	E
131	S	450	L
132	E	480	T
187	T	496	T
279	E	515	I
286	R	522	S
339	S	523	E
347	E	524	Y

The chances that any one position might have a fixed point may be 65%, but the chances that all of these positions have fixed points is about 0.003%, and that positions 51, 132, 279, 347, 429, and 523 all have the *same* fixed point is so much smaller. It looks like we will be able to narrow down the possibilities to one or a few.

We have 93,750 settings for the twenties switches to check, so let's begin with the first one: left L = middle M = right R = 1 and fast = left, medium = middle, slow = right. For these settings and what we know about the sixes switch, here are the overall permutations of the three twenties switches at the positions of interest in the text:

position in text	overall permutation of twenties switches																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
17	12	5	19	1	8	9	18	10	6	4	20	2	13	16	17	11	7	15	14	3
51	5	13	19	16	7	1	20	6	4	18	15	11	14	8	3	10	9	17	12	2
68	2	19	14	16	1	15	17	18	5	9	6	10	20	3	4	12	8	7	13	11
80	4	9	13	19	15	2	5	14	1	20	6	3	7	17	10	11	18	12	8	16
85	18	15	17	14	8	13	7	5	16	12	19	10	20	6	11	2	1	3	9	4
131	2	20	8	3	10	16	13	17	1	5	9	7	4	6	14	18	19	11	12	15
132	10	1	9	20	13	19	7	5	16	11	2	14	18	4	6	8	3	17	15	12
187	6	20	8	14	19	10	7	2	3	12	11	13	15	9	1	17	4	16	5	18
279	7	10	13	8	1	15	6	2	20	11	17	5	19	14	3	18	12	16	4	9

286	5	14	17	18	16	19	11	7	15	9	13	12	4	20	1	10	3	8	6	2
339	14	3	1	2	7	4	9	13	6	11	15	8	10	16	5	17	18	12	20	19
347	10	20	18	19	4	7	8	9	2	6	11	5	15	12	14	1	17	13	16	3
385	9	3	18	1	17	6	14	13	5	8	10	20	19	4	7	11	16	12	2	15
388	20	19	7	11	5	14	4	15	16	3	6	9	10	1	17	12	18	2	8	13
420	9	15	1	3	14	6	8	12	7	11	20	17	5	10	13	18	19	4	16	2
426	8	14	6	12	17	11	19	10	3	16	4	1	20	15	13	9	2	7	5	18
429	11	18	19	4	8	13	17	16	15	2	12	20	9	7	1	6	10	14	3	5
450	10	20	6	7	11	2	19	4	3	16	15	17	5	9	12	13	8	1	18	14
480	13	11	9	16	4	2	12	20	14	8	17	5	1	3	7	6	10	19	15	18
496	4	6	7	15	19	9	1	18	20	8	17	12	5	2	11	3	16	10	14	13
515	13	9	17	10	16	1	7	5	15	2	14	3	11	20	18	12	4	8	6	19
522	17	9	14	6	5	11	15	19	3	2	4	18	1	20	13	12	10	8	16	7
523	5	3	4	9	15	10	18	2	11	8	17	13	12	1	20	14	6	19	7	16
524	1	14	6	10	18	20	9	7	4	16	2	19	11	5	3	17	8	15	13	12

Position 17 has a fixed point (13), but the fixed point at 385 (6) does not match. Nine of these permutations don't even have fixed points. So these settings are inconsistent with our plaintext-ciphertext pair. After searching all 93,750 possibilities, we find four settings that give fixed points at all of the relevant twenty-four positions in the text:

left *L*	middle *M*	right *R*	fast	medium	slow	fixed point(s) at 17	fixed point(s) at 385
20	15	2	middle	left	right	4	6
8	18	25	middle	right	left	7	14
21	4	16	middle	right	left	2	2, 10
22	11	15	right	middle	left	12	2

Only one of these has the same fixed point at positions 17 and 385, which is what we need. So that must be the correct key.

Can we now find the plugboard settings for the twenties? Here are the fixed points for sixes = 8, left = 21, middle = 4, right = 16, fast = middle, medium = right, slow = left:

position in text	overall permutation of twenties switches 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
17	17	2	14	8	1	18	20	11	4	12	15	10	6	5	13	7	19	9	16	3
51	13	6	5	18	3	16	15	11	17	9	12	7	8	19	2	4	1	10	14	20
68	16	6	17	10	14	3	11	8	19	18	5	2	7	4	15	12	9	1	20	13
80	20	15	9	16	2	14	6	12	5	8	11	4	13	7	18	1	17	10	3	19
85	3	18	5	8	20	6	4	12	14	11	2	10	17	1	15	13	7	19	16	9
131	13	1	6	8	10	20	16	4	18	14	11	12	7	9	2	17	3	5	15	19
132	12	7	5	14	11	4	10	17	1	13	3	8	18	15	16	6	9	2	19	20
187	19	14	8	11	13	12	7	20	2	5	3	9	16	15	4	18	17	10	6	1

279	10	15	9	17	2	5	8	11	6	7	16	18	19	14	1	4	3	12	13	20
286	8	19	5	4	12	6	2	20	15	1	13	7	9	11	3	16	10	18	17	14
339	3	15	17	20	14	10	16	9	6	7	11	4	8	5	2	1	19	13	18	12
347	8	10	19	15	16	18	13	7	5	12	6	1	4	9	3	14	2	11	17	20
385	5	2	1	16	13	3	4	19	17	10	8	15	12	7	9	6	18	20	14	11
388	17	6	3	7	12	15	16	9	20	11	14	19	10	8	5	4	2	18	13	1
420	12	8	17	5	1	11	16	15	4	3	20	13	14	7	10	19	9	18	2	6
426	6	17	5	15	11	4	8	7	9	1	2	19	13	18	12	16	3	14	20	10
429	18	11	14	12	15	1	7	13	5	2	17	4	6	8	16	3	19	9	10	20
450	17	8	10	16	1	11	18	12	6	19	3	20	5	9	15	2	13	4	7	14
480	4	7	1	5	20	19	3	13	16	8	6	12	11	2	9	10	17	15	18	14
496	11	10	4	16	8	9	7	13	14	20	6	19	1	18	3	12	17	5	15	2
515	2	19	17	4	9	6	20	7	18	13	15	8	10	11	5	16	3	12	1	14
522	2	20	9	6	12	8	1	17	19	13	11	5	18	10	15	14	4	7	3	16
523	3	10	6	12	17	7	4	9	15	1	16	14	13	2	8	18	19	5	11	20
524	1	7	8	17	5	3	18	4	20	10	15	19	9	16	2	6	11	12	13	14

By comparing to the letters that are enciphered to themselves, we find

twenties letter	twenties wire	plugboard letter
B	1	
C	2	N
D	3	
F	4	
G	5	
H	6	
J	7	
K	8	F
L	9	
M	10	
N	11	S
P	12	
Q	13	
R	14	
S	15	L
T	16	
V	17	T
W	18	R
X	19	
Z	20	E

We cannot place the I, W, and Y just yet, because there are too many choices for each of them. To fill in the rest of this table, we need to locate each of the missing letters in the plaintext, check to see that we already know the mapping of the corresponding ciphertext letter, find the overall twenties permutation

at that point, and see on which line the plaintext letter must be. For example, at position 66 in the text, B is enciphered to F, which we know to be on wire 8. The overall twenties permutation at that point is

(18 19 11 13 4 9 8 15 5 3 1 2 7 10 12 16 20 6 17 14)

Since 8 is in the 7th place in the permutation, B must be on wire 7. We play the same game for the remaining twenties letters.

plaintext letter	ciphertext letter	position in text	overall permutation of twenties switches 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
B	F on 8	66	18 19 11 13 4 9 8 15 5 3 1 2 7 10 12 16 20 6 17 14
G	L on 15	307	18 9 13 12 1 19 11 5 7 17 20 2 8 4 10 15 14 16 6 3
H	R on 18	15	6 9 12 11 17 5 7 16 8 14 10 3 18 13 1 19 2 20 4 15
I	N on 2	47	5 11 13 2 19 12 18 4 10 20 8 15 9 3 17 6 14 7 1 16
J	none		
P	F on 8	208	12 2 1 18 13 17 10 3 6 16 11 19 9 8 20 14 4 7 5 15
Q	none		
U	N on 2	40	6 4 17 18 14 16 11 7 3 2 12 19 8 20 13 10 1 15 5 9
V	E on 20	96	3 11 8 5 19 10 16 1 9 18 2 6 13 7 12 17 14 4 20 15
W	R on 18	14	7 11 20 3 10 19 4 2 18 16 6 12 1 17 14 15 13 5 9 8
X	none		
Y	L on 15	29	19 12 10 11 15 18 20 6 8 4 17 16 2 13 14 5 7 1 9 3
Z	none		

We can now fill in more of the list. For example, B must be on wire 7, since that one is mapped to 8 at position 66.

twenties letter	twenties wire	plugboard letter
B	1	
C	2	N
D	3	
F	4	I
G	5	Y
H	6	
J	7	B
K	8	F
L	9	W
M	10	U
N	11	S
P	12	
Q	13	H
R	14	P
S	15	L
T	16	G

V	17	T
W	18	R
X	19	V
Z	20	E

The twenties plugboard setting is therefore _N_IY_BFWUS_HPLGTRVE, where the blanks are for letters that do not appear in the plaintext and are therefore irrelevant. We can find them, however, because they appear in the ciphertext.

plaintext letter	ciphertext letter	position in text	overall permutation of twenties switches 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
E on 20	J	367	18 17 12 2 5 19 16 8 14 4 6 10 7 11 1 13 15 20 9 3
E on 20	Q	16	12 17 7 5 6 13 11 9 19 18 2 20 16 14 10 8 4 3 15 1
E on 20	X	11	9 16 19 15 2 13 8 5 14 10 11 6 7 4 18 1 17 20 3 12
E on 20	Z	147	4 15 16 11 8 19 12 14 10 17 1 3 9 2 13 18 5 7 20 6

And now we can fill in the remaining spaces:

twenties letter	twenties wire	plugboard letter
B	1	Q
C	2	N
D	3	J
F	4	I
G	5	Y
H	6	Z
J	7	B
K	8	F
L	9	W
M	10	U
N	11	S
P	12	X
Q	13	H
R	14	P
S	15	L
T	16	G
V	17	T
W	18	R
X	19	V
Z	20	E

We now have the entire key: $S = 8$, $L = 21$, $M = 4$, $R = 16$, fast = M, medium = R, slow = L, sixes = DCMAKO, twenties = QNJIYZBFWUSXHPLGTRVE.

Reading and references

Barjol Lami, Gledis Kallco, Nicholas Guo, and Sean Shi; “Cryptanalysis of Purple, Japanese WWII Cipher Machine,” 2019, https://courses.csail.mit.edu/6.857/2019/project/24-Lami-Kallco-Guo-Shi.pdf

Programming tasks

1. Automate as much of the process as you like.

Exercises

1. Find as much of the key as you can from this pair of texts, taken from two parts of *The Purple Cloud* by M. P. Shiel.

```
ALONETHATSAMEDAYIBEGANMYWAYSOUTHWARDANDFORFIVEDAYSMADEGOODPROGRE
SSONTHEEIGHTHDAYINOTICEDSTRETCHEDRIGHTACROSSTHESOUTHEASTERNHORIZ
ONAREGIONOFPURPLEVAPOURWHICHLURIDLYOBSCUREDTHEFACEOFTHESUNANDDAY
AFTERDAYISAWITSTEADILYBROODINGTHEREBUTWHATITCOULDBEIDIDNOTUNDERS
TANDITSNATUREITSORIGINREMAINSOFCOURSENOTHINGBUTMATTEROFCONJECTUR
EFORITLEAVESNOLIVINGTHINGBEHINDITNORGODKNOWSISTHATOFANYMOMENTNOW
TOUSWHOREMAINTHERUMOURTHATITISASSOCIATEDWITHANODOUROFALMONDSISDE
CLAREDONHIGHAUTHORITYTOBEIMPROBABLEBUTTHEMOROSEPURPLEOFITSIMPEND
INGGLOOMHASBEENATTESTEDBYTARDYFUGITIVESFROMTHEFACEOFITSROLLINGAN
DSMOKYMARCH

KUPCLBXBBVEPQNGYODCKZXWYYSIJYELAYJFSLRGKWJAOUALAYNINUKQOICOCYNZN
ZDIZJNHDYTSXUKVOYNWQWQZXVTKGBDZALRWAQQKAVONZHTZAWRQZSJLZVJHCOHMZ
YXLCNCWYHWDICEIBLFNMIGTPCWAXBCQWNQIYZHZEGJXBTCFZGXYEHXLXXZSCTKJP
XVBXGUVPOVCOMRBGNQSOUONRIPAWZUSUSVVVCHMDFAWFQWXNEEGOLIXSIXHEVURE
QZVLMNUHHKGEQWRVODIHMZGZWRPSCYATYCUFLVOSAYALHLDPHCDUJMLBWZNTSDEV
CGOEIQEVEXECSPJMHPZHGQYDZREAWGUMQNWZUWSDCIMSYKTRHFMFNJPIWOFUAZWM
RYKSWFWJFIEIGBVXEGOYHEDEBAMTPFLBZPVWAQUNYOLSSDOKWCSYNERPWRLRWDRN
AAQRGEWBLPJDCUJEMROLWCMXRWMYSPEFSXDQAVNLZPWUPFAPENPKJWFWJDOOOAEH
PTGXFPMPAQENNBKZVHTDBFVZMJKNHPHTRICYBEEVUYYFZBALLZWHYSJZMUAPXLFE
HNPWGYOSQTJ
```