DES (Data Encryption Standard) (January 15th, 1977).

- $\mathcal{P} = \mathcal{C} = \{0,1\}^{64}$.
- $\mathcal{K} = \{0,1\}^{56}$.
- Encoding bit-string x with the key K:
  1. Let $x_0 = IP(x)$, where IP is a certain permutation of bits. Let L_0 [R_0] be the first [last] 32 bits of x.
  2. 16 rounds of Feistel construction:

$$L_i = R_{i-1} \qquad R_i = L_{i-1} \oplus f(R_{i-1}, K_i)$$

     Here $1 \leqslant i \leqslant 16$, $K_i \in \{0,1\}^{48}$ consist of certain 48 bits of K.
  3. Let $y = IP^{-1}(R_{16}L_{16})$. y is the ciphertext.

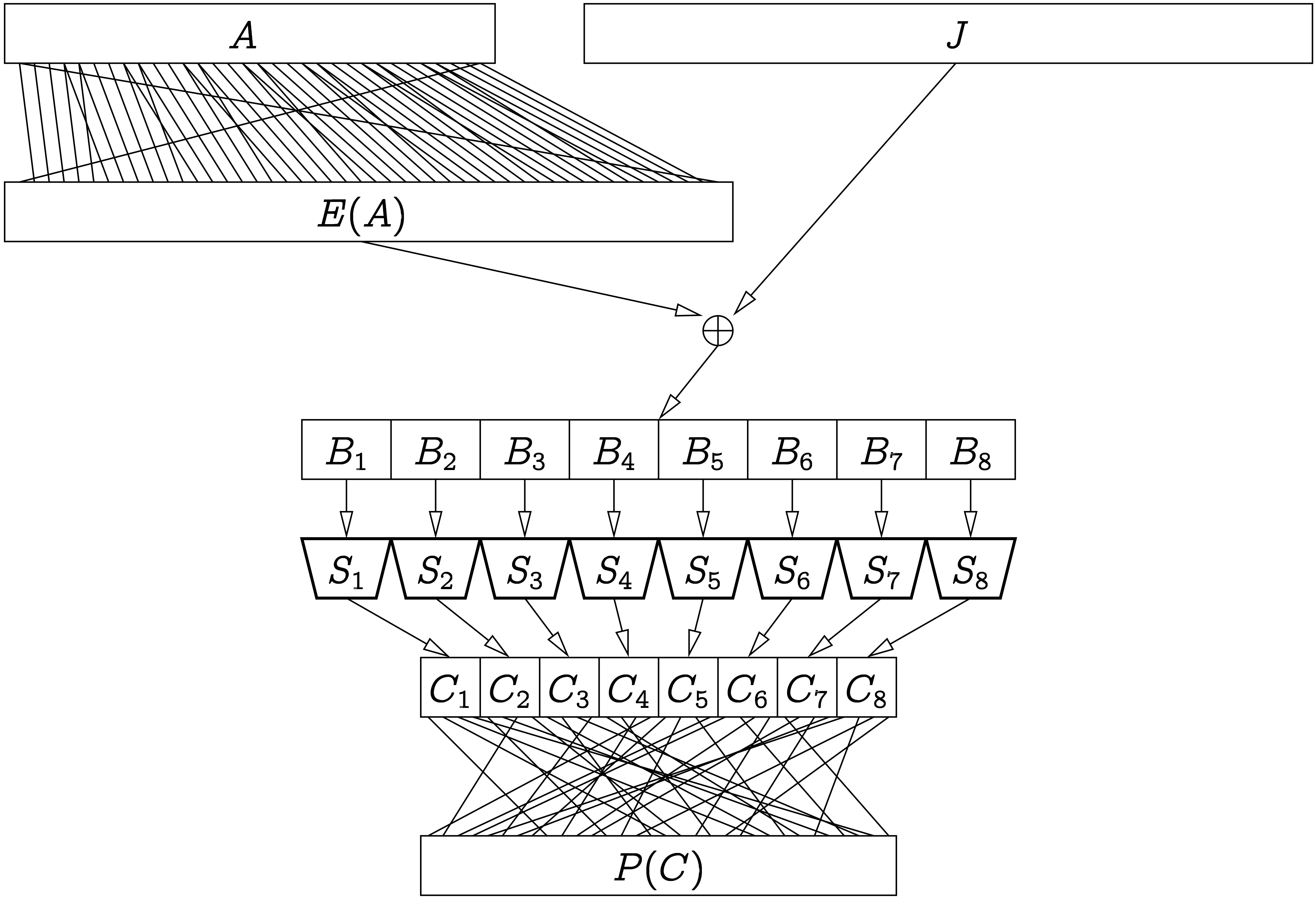
$f : \{0,1\}^{32} \times \{0,1\}^{48} \to \{0,1\}^{32}$. $f(A, J)$ works as follows:

1. "Expand" A to $E(A)$ of length 48. The function E outputs the bits of its argument in certain order (16 bit positions occur once and 16 occur twice).

2. Let $B_1 \cdots B_8 = E(A) \oplus J$, where $B_i \in \{0,1\}^6$.

3. Let $C_i = S_i(B_i)$, where $S_i : \{0,1\}^6 \to \{0,1\}^4$ is a fixed mapping. (the *S-box*)

4. return $P(C_1 \cdots C_8)$ where P is a certain permutation of bits.



A
J
$E(A)$
B_1 B_2 B_3 B_4 B_5 B_6 B_7 B_8
S_1 S_2 S_3 S_4 S_5 S_6 S_7 S_8
C_1 C_2 C_3 C_4 C_5 C_6 C_7 C_8
$P(C)$

Decryption: like encryption, but round keys taken in order $K_{16}, K_{15}, \ldots, K_1$.

In the standard, the encryption key is actually 8 bytes long.

- The least significant bit in each byte is a parity check bit. Not used in actual encryption.
- The number of 1-s in each byte is odd.

Differential cryptanalysis — a chosen-plaintext attack.

For reduced-round DES, it is more efficient than brute-force search.

n-round DES — $L_0 R_0 \mapsto L_n R_n$. We ignore the bit-permutations IP, IP^{-1}, nor do we swap L_n and R_n.

Idea, given two bit-strings $L_0 R_0$ and $L_0^* R_0^*$ with a fixed xor $L_0' R_0' = L_0 R_0 \oplus L_0^* R_0^*$, we compare the xor-s of their encryptions. This will help us to exclude certain values for the key.

We attempt to reconstruct the xor-s of the intermediate computations.

Let $B' \in \{0,1\}^6$ and $1 \leqslant j \leqslant 8$. For all $B \in \{0,1\}^6$ consider the value $S_j(B) \oplus S_j(B \oplus B')$.

- The pairs $(B, B \oplus B')$ range over all possible pairs of six-bit strings with xor B'.
- The bit-strings $S_j(B) \oplus S_j(B \oplus B')$ range over four-bit strings.
  - Typically, not all four-bit strings are achieved.
  - If the output xor of an S-box is C' then certain input xor-s are excluded.

For $B' \in \{0,1\}^6$, $C' \in \{0,1\}^4$ and $j \in \{1,\ldots,8\}$ define

$$\begin{aligned} IN_j(B',C') &= \{B \in \{0,1\}^6 \mid S_j(B) \oplus S_j(B \oplus B') = C'\} \\ N_j(B',C') &= |IN_j(B',C')| \end{aligned}$$

The sets $IN_j(B',C')$ can be computed from the definition of S-boxes. There are 8192 such sets — not too many.

About a fifth of the sets $IN_j(B',C')$ is empty.

Let now $B, B^* \in \{0,1\}^{48}$ be two inputs to (all) S-boxes in a computation of f with $B' = B \oplus B^*$. Then

$$B' = B \oplus B^* = E(A) \oplus J \oplus E(A^*) \oplus J = E(A) \oplus E(A^*)$$

Denote $E(A)$ by E and $E(A^*)$ by E^*. We see that B' does not depend on J. If $C = S(B)$ and $C^* = S(B^*)$ then $C' = C \oplus C^*$ depends on J.

Let

$$test_j(E_j, E_j^*, C_j') = \{B_j \oplus E_j \mid B_j \in IN_j(E_j', C_j')\}$$

where $E_j, E_j^* \in \{0,1\}^6$, $C_j' \in \{0,1\}^4$ and $E_j' = E_j \oplus E_j^*$.

Theorem. Let E_j, E_j^* be two inputs to the S-box S_j (before being xor-ed with the key bits J_j). Let C_j' be the output xor of these inputs. Then $J_j \in test_j(E_j, E_j^*, C_j')$.

To obtain a unique result, use several triples E, E^*, C'.

Example: three-round DES. If the plaintext is $L_0 R_0$ and ciphertext is $L_3 R_3$ then

$$R_3 = L_2 \oplus f(R_2, K_3) = L_0 \oplus f(R_0, K_1) \oplus f(R_2, K_3)$$
$$L_3 = R_2 = L_1 \oplus f(R_1, K_2) = R_0 \oplus f(R_1, K_2)$$

Pick another plaintext $L_0^* R_0^*$. Then $R_3' = R_3 \oplus R_3^*$ equals

$$R_3' = L_0' \oplus f(R_0, K_1) \oplus f(R_0^*, K_1) \oplus f(R_2, K_3) \oplus f(R_2^*, K_3)$$

We choose $R_0^* = R_0$. Then $R_0' = \mathbf{0}^{32}$ and

$$R_3' = L_0' \oplus f(R_2, K_3) \oplus f(R_2^*, K_3) \ .$$

We know L_0' and R_3'. Hence we can compute

$$f(R_2, K_3) \oplus f(R_2^*, K_3) = R_3' \oplus L_0' \quad .$$

$f(R_2, K_3) = P(C)$ and $f(R_2^*, K_3) = P(C^*)$ for some S-box outputs C and C^*. We have $C' = C \oplus C^* = P^{-1}(R_3' \oplus L_0')$.

We know $R_2 = L_3$ and $R_2^* = L_3^*$. The inputs to the S-box are $E(R_2) \oplus K_3$ and $E(R_2^*) \oplus K_3$.

We know E, E^*, C' for the third round. We can compute the sets $test_1, \ldots, test_8$ and construct candidate round keys K_3.

Using several such triples E, E^*, C' we narrow down the set of candidate round keys K_3.

A one-round characteristic is a quantity

$$L_0'R_0' \xrightarrow{p_1} L_1'R_1'$$

where

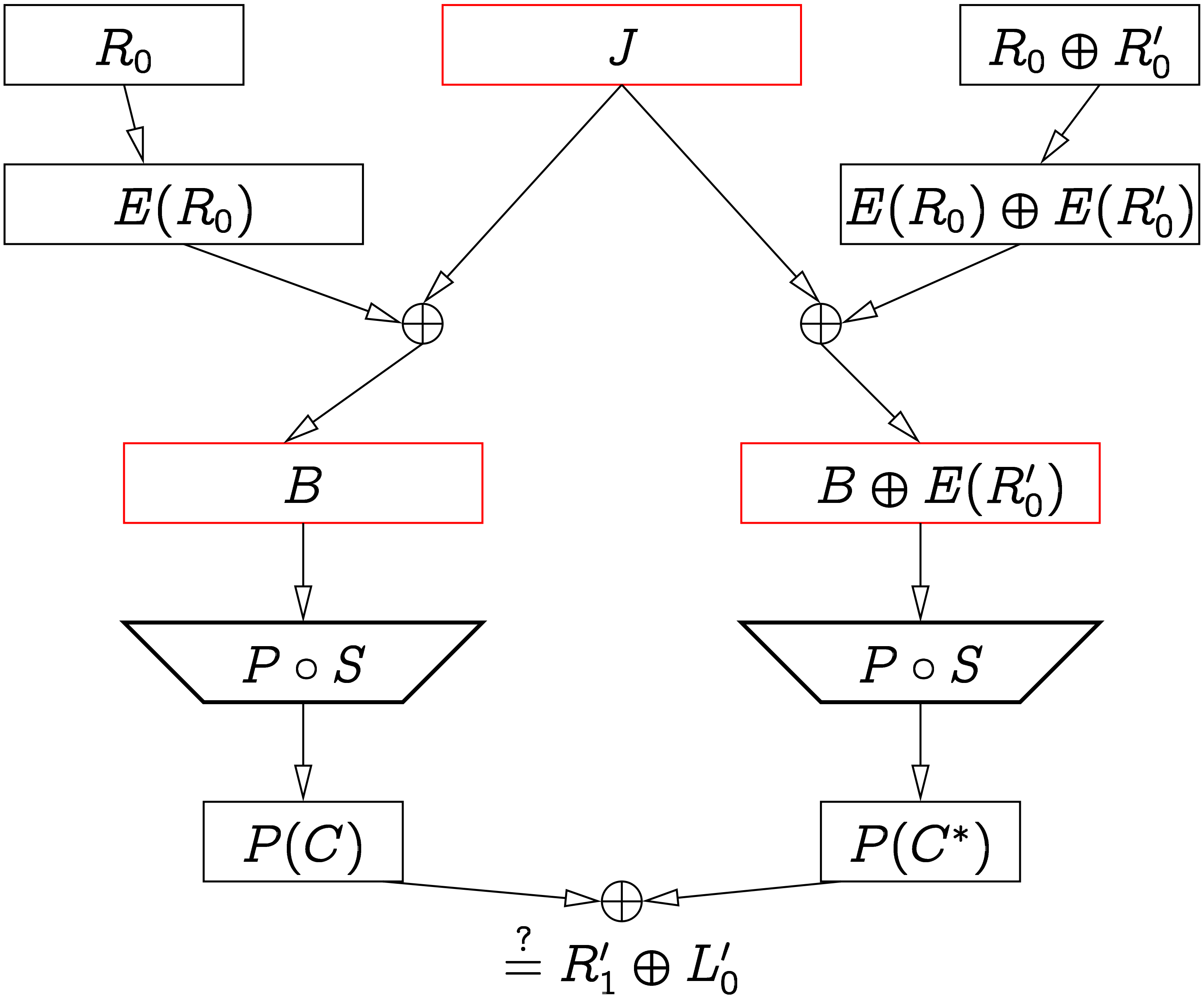
- $L_1' = R_0'$;
- For any choice of L_0, R_0, the quantity p_1 is the probability that (taken over uniformly chosen $J \in \{0,1\}^{48}$)

$$\big(L_0 \oplus f(R_0, J)\big) \oplus \big((L_0 \oplus L_0') \oplus f(R_0 \oplus R_0', J)\big) = R_1'$$

  or that

$$f(R_0, J) \oplus f(R_0 \oplus R_0', J) = R_1' \oplus L_0' \ .$$

That probability does not depend on R_0 either.



R_0
J
$R_0 \oplus R_0'$
$E(R_0)$
$E(R_0) \oplus E(R_0')$
B
$B \oplus E(R_0')$
$P \circ S$
$P \circ S$
$P(C)$
$P(C^*)$
$\stackrel{?}{=} R_1' \oplus L_0'$

p_1 is the probability that

$$S(B) \oplus S(B \oplus E(R_0')) = P^{-1}(R_1') \oplus P^{-1}(L_0')$$

where $B \in \{0,1\}^{48}$ has been uniformly chosen.

An n-round characteristic is

$$L_0'R_0' \xrightarrow{p_1} L_1'R_1' \xrightarrow{p_2} \cdots \xrightarrow{p_n} L_n'R_n'$$

where each $L_{i-1}'R_{i-1}' \xrightarrow{p_i} L_i'R_i'$ is a one-round characteristic.

The probability of such a characteristic is $p_1 \cdots p_n$.

Some one-round characteristics:

$$xxxxxxxx_{16}|00000000_{16} \xrightarrow{1} 00000000_{16}|xxxxxxxx_{16}$$

$$00000000_{16}|60000000_{16} \xrightarrow{14/64} 60000000_{16}|00808200_{16}$$

Second example: $E(R'_0) = 001100\cdots0_2$. Hence the inputs to S-boxes $S_2, \ldots, S_8$ are equal, but the inputs to S_1 differ by 001100.

The probability that the outputs to S_1 differ by $x \in \{0,1\}^4$ is $N_1(001100_2, x)/64$. In particular, $N_1(001100_2, 1110_2) = 14$.

The output difference of S-boxes is $111000\cdots0_2$ with probability $14/64$. The bit-permutation P brings those three 1-s to the positions shown above.

Example: six-round DES.

$$R_6 = R_4 \oplus f(R_5, K_6) = L_3 \oplus f(R_3, K_4) \oplus f(R_5, K_6)$$

$$R_6' = L_3' \oplus f(R_3, K_4) \oplus f(R_3^*, K_4) \oplus f(R_5, K_6) \oplus f(R_5^*, K_6)$$

We try to find K_6.

A three-round characteristic:

$$40080000_{16}|04000000_{16} \xrightarrow{1/4} 04000000_{16}|00000000_{16} \xrightarrow{1}$$
$$00000000_{16}|04000000_{16} \xrightarrow{1/4} 04000000_{16}|40080000_{16}$$

If $L_0'R_0' = 40080000_{16}|04000000_{16}$ then
$L_3'R_3' = 04000000_{16}|40080000_{16}$ with probability $1/16$.

Assume that this happens, i.e. we know L_3' and R_3'. We also know R_6' and $R_5' = L_6'$.

$E(R_3') = 001000|000000|000001|010000|0\cdots0$. I.e. the input (and also output) xor-s to S_2, S_5, S_6, S_7, S_8 in the fourth round are zero. We try to find the corresponding 30 bits of K_6.

$$R_6' = L_3' \oplus f(R_3, K_4) \oplus f(R_3^*, K_4) \oplus f(R_5, K_6) \oplus f(R_5^*, K_6)$$

and certain 20 bits of $f(R_3, K_4)$ and $f(R_3^*, K_4)$ are equal. These 20 bits in $f(R_5, K_6) \oplus f(R_5^*, K_6)$ are equal to the same bits in R_6'.

We know the output xor-s of S_2, S_5, S_6, S_7, S_8 in the sixth round. We also know the inputs to these S-boxes (as we know $R_5 = L_6$ and $R_5^* = L_6^*$).

We know the triples E_i, E_i^*, C_i' for the sixth round, where $i \in \{2, 5, 6, 7, 8\}$. We can compute the sets $test_i$ and find the candidate keys.

We also get noise (because our certainty in $L_3' R_3'$ was only $1/16$), but the right key should stick out.

To find the right key more quickly:

We have the plaintext pairs $(x_1, x_1^*), \ldots, (x_N, x_N^*)$ with $x_i \oplus x_i^* = L_0' R_0'$.

Each of these pairs defines a quintuple of sets $(test_2^{(i)}, test_5^{(i)}, test_6^{(i)}, test_7^{(i)}, test_8^{(i)})$.

For each i: if this quintuple of sets contains the empty set, then discard it.

A set $\{i_1, \ldots, i_n\} \subseteq \{1, \ldots, N\}$ is allowable if

$$\bigcap_{k=1}^{n} test_j^{(i_k)} \neq \emptyset \text{ for all } j \in \{2, 5, 6, 7, 8\} \ .$$

We search for an allowable set of maximum cardinality (using backtracking).

We have found 30 bits of the key. The characteristic

$$00200008_{16}|00000400_{16} \xrightarrow{1/4} 00000400_{16}|00000000_{16} \xrightarrow{1}$$
$$00000000_{16}|00000400_{16} \xrightarrow{1/4} 00000400_{16}|00200008_{16}$$

allows us to find further 12 (those corresponding to the inputs of S_1 and S_4). The remaining 14 bits can be brute-forced.

A two-round characteristic:

$$19600000_{16}|00000000_{16} \xrightarrow{1} 00000000_{16}|19600000_{16}$$
$$\xrightarrow{14\cdot 8\cdot 10/(64)^3} 19600000_{16}|00000000_{16}$$

The second fraction is about $1/234$. Iterating this characteristic 6.5 times gives a 13-round characteristic of probability $1/234^6$. This is the best-known characteristic for cryptanalysing full 16-round DES.