A4. Совместимость фильтров Блума

Вопрос 1

Верно ли, что F(AB) будет выдавать положительные ответы о принадлежности объектов из $A \cap B$?

Да, верно.

Пусть $x \in A \cap B$. Тогда x был вставлен и в F(A), и в F(B). Значит для каждой хеш-функции h_i бит на позиции h_i(x) установлен в 1 и в F(A), и в F(B).

При побитовом И: если бит равен 1 в обоих фильтрах, он будет равен 1 и в F(AB). Следовательно, все биты, необходимые для подтверждения принадлежности x, будут установлены в F(AB).

Поэтому для любого $x \in A \cap B$ фильтр F(AB) ответит «да» (положительный ответ).

Стоит отметить, что F(AB) может также давать ложноположительные ответы для элементов, не принадлежащих $A \cap B$ — это стандартное свойство фильтров Блума.

Вопрос 2

Верно ли, что F(AB) будет в точности соответствовать другому фильтру, построенному последовательной вставкой объектов из $A \cap B$?

Нет, не верно.

Обозначим фильтр, построенный вставкой элементов $A \cap B$, как $F(A \cap B)$.

F(AB) = F(A) AND F(B) может содержать «лишние» единичные биты, которых не будет в $F(A \cap B)$.

Пример: пусть элемент $a \in A \setminus B$ устанавливает бит на позиции 5 в F(A), а элемент $b \in B \setminus A$ устанавливает бит на позиции 5 в F(B). Тогда в F(AB) бит 5 будет равен 1 (он равен 1 в обоих фильтрах). Но ни a, ни b не принадлежат $A \cap B$, поэтому если бит 5 не нужен ни одному элементу из $A \cap B$, в $F(A \cap B)$ он будет равен 0.

Таким образом, $F(AB) \neq F(A \cap B)$ в общем случае. Гарантируется только, что $F(A \cap B)$ является «подмножеством» F(AB) побитово, то есть каждый единичный бит в $F(A \cap B)$ будет единичным и в F(AB), но не наоборот.