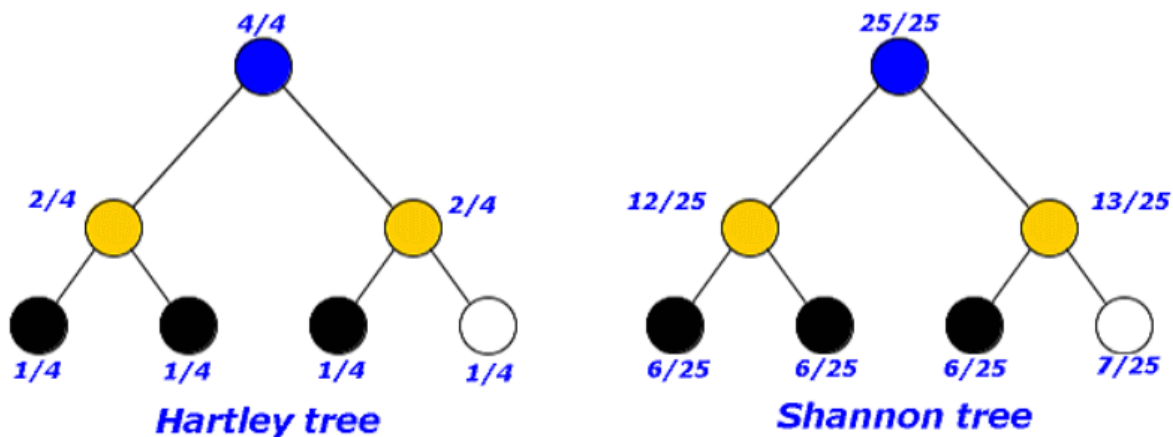


Задача о шарах



Q. В корзине лежат черные и белые шары, причем черных - 18. Сообщение о том, что достали белый шар, несет 2 бита информации. Сколько всего шаров в корзине?

A. Это означает, что за два вопроса да/нет (два шага по дереву выбора) мы расслаиваем пространство состояний равными долями на белые/черные. То есть, белым принадлежит $1/4$ всего пространства состояний и черным - оставшиеся $3/4$. Три четверти - это 18, следовательно $1/4 \rightarrow 18/3 = 6$. Белых шаров 6, всего шаров - 24.

Q. Давайте так - у нас 18 черных и 7 белых шаров, сколько битов информации будет нести сообщение "из корзины достали белый шар"?

$$I = \log_2((18+7)/7) = 1,836501268$$

Будет ли такая информация нести 2 бита информации? Можно ли ее получить за два вопроса, предполагающих только ответы Да/Нет?

Чем ситуация с 7 белыми и 18 черными шарами будет отличаться от ситуации с 6 белыми и 18 черными шарами?

A. Начну с конца:

Чем ситуация с 7 белыми и 18 черными шарами будет отличаться от ситуации с 6 белыми и 18 черными шарами?

Условие исходной задачи сформулировано так, что дерево выбора сбалансировано и к нему применима формула Хартли (пространство событий может быть поделено на четыре равных области).

В новой формулировке это уже невозможно и требуется использование более общей формулы Шеннона.

У нас 18 черных и 7 белых шаров, сколько битов информации будет нести сообщение "из корзины достали белый шар"?

Считая по Шеннону, чуть более полубита:

$$-(7/25) * \log_2(7/25) = -0.280 * -1.83650126771712 = 0.514220354960794$$

Для наглядности, посчитаем шенноновскую энтропию в обоих вариантах (в силу асимметрии выбора, она всегда меньше единицы).

Для исходной задачи (18 + 6 = 24):

$$18/24 = 0.750$$

$$6/24 = 0.250$$

$$\log_2(18/24) = -0.415037499278844$$

$$\log_2(6/24) = -2$$

$$-(18/24) * \log_2(18/24) = -0.750 * -0.415037499278844 = 0.311278124459133$$

$$-(6/24) * \log_2(6/24) = -0.250 * -2 = 0.5$$

$$\text{Sum: } 0.811278124459133$$

Для нового варианта (18 + 7 = 25):

$$18/25 = 0.720$$

$$7/25 = 0.280$$

$$\log_2(18/25) = -0.473931188332412$$

$$\log_2(7/25) = -1.83650126771712$$

$$-(18/25) * \log_2(18/25) = -0.720 * -0.473931188332412 = 0.341230455599337$$

$$-(7/25) * \log_2(7/25) = -0.280 * -1.83650126771712 = 0.514220354960794$$

$$\text{Sum: } 0.855450810560131$$

Можно ли ее получить за два вопроса, предполагающих только ответы Да/Нет?

На этот вопрос невозможно ответить в рамках парадигмы равновероятных ответов Да/Нет и неделимых шаров. Невозможно нацело разделить нечетное число на $2^2 = 4$. Иными словами, кратность 3:1 - необходимое условие для построения двухуровневого сбалансированного дерева выбора.

С другой стороны, если пространство состояний разрешено делить на неравные части, то при двух возможных вариантах ответа: "черный-белый" и в первом и во втором случае случае нужный цвет может быть найден за один вопрос - достаточно

ткнуть в любую и спросить, какого она цвета. И если цвет не тот, значит нужный - в другой.