Definition: Prefix-freeness

One convenient way to guarantee that a code is unique decodable is to require it to be prefix-free:

Definition: Prefix-free code

A binary symbol code $C : \mathcal{X} \to \{0,1\}^*$ is prefix-free (or: **instantaneous**) if for all $x, x' \in \mathcal{X}$ with $x \neq x'$, $C(x)$ is *not* a prefix of $C(x')$.

With a prefix-free encoding, the elements $x_1, \ldots, x_m$ can be uniquely recovered from $C(x_1)|\cdots|C(x_m)$, simply by reading the encoding from left to right one bit at a time: by prefix-freeness it will remain unambiguous as reading continues when the current word terminates and the next begins. This is a loose argument for the following proposition:

Proposition

If a code $\mathcal{C}$ is prefix-free and $\mathcal{C} \neq \perp$ then $\mathcal{C}$ is uniquely decodable.

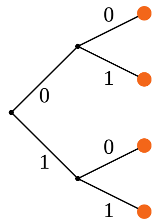
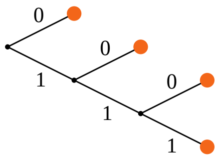
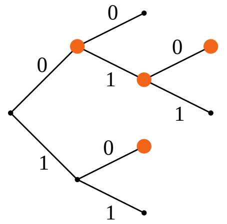
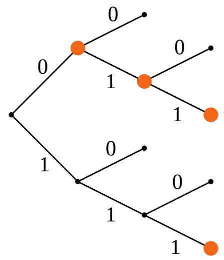
The other direction does not hold: uniquely decodable codes need not be prefix-free. A prefix-free code is appealing from an efficiency point of view, as it allows to decode "on the fly". For a general uniquely decodable code one may possibly have to inspect all bits in the entire string before being able to even recover the first word.

Example

The following are all codes for the source P_X, with $\mathcal{X} = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$:

x	$P_X(x)$	$C_1(x)$	$C_2(x)$	$C_3(x)$	$C_4(x)$
$\mathtt{a}$	0.5	00	0	0	0
$\mathtt{b}$	0.25	01	10	010	01
$\mathtt{c}$	0.125	10	110	01	011
$\mathtt{d}$	0.125	11	111	10	111

These codes can be visualised a as binary trees, with marked codewords, as follows:

			
C_1	C_2	C_3	C_4

Which of these codes are uniquely decodable? Which are prefix-free?

Show solution

C_1 and C_2 are prefix-free, and therefore also uniquely decodable. C_3 is not uniquely decodable, as $C_3(ad) = C_3(b)$. C_4 is not prefix-free, but it is uniquely decodable, since it can be decoded from right to left (it is ``suffix-free"). Note that the binary trees for the prefix-free codes C_1, C_2 only have codewords at the leaves.