13 Imperfection Ratio

Theorem 13.1 (Gerke, McDiarmid)

$$\mathrm{imp}(G) = \max\{\chi^*(G,x) : x \in \mathrm{QSTAB}(G)\} =$$
$$\min\{t : \mathrm{QSTAB}(G) \subseteq t\,\mathrm{STAB}(G)\} =$$
$$\max\{x \cdot y : x \in \mathrm{QSTAB}(G), y \in \mathrm{QSTAB}(\overline{G})\}$$

Proof.

Claim (1) $\omega(G,x) \leq t$ if and only if $x \in t\,\mathrm{QSTAB}(G)$.

By the scaling property, $\omega(G,x) \leq t$ is equivalent to $\omega(G, \frac{1}{t}x) \leq 1$, which is clearly equivalent to $\frac{1}{t}x \in \mathrm{QSTAB}(G)$, and so the claim follows.

Claim (2) $\chi^*(G,x) \leq t$ if and only if $x \in t\,\mathrm{STAB}(G)$.

By definition, $\chi^*(G,x) \leq t$ implies that there are stable sets $S_1, \ldots, S_k$ and constants $y_1, \ldots, y_k$ such that $y_i \geq 0$, $\sum_{i:v \in S_i} y_i \geq x_v$ for any $v \in V(G)$ and $y_1 + \cdots + y_k \leq t$. Thus $\frac{1}{t}y_1 + \cdots + \frac{1}{t}y_k \leq 1$ and hence the vector

$$z = \frac{1}{t}y_1 \mathbf{1}_{S_1} + \cdots + \frac{1}{t}y_k \mathbf{1}_{S_k}$$

is a convex combination of incidence vectors of stable sets, and so $z \in \mathrm{STAB}(G)$. But $\frac{1}{t}x \leq z$ coordinatewise, and so the "only if" direction follows.

One can easily reverse the argument above to get the "if" direction.

Claim (3) $\mathrm{imp}(G) \leq t$ if and only if $\chi^*(G,x) \leq t$ for any $x \in \mathrm{QSTAB}(G)$.

$\mathrm{imp}(G) \leq t$ is equivalent to the following:

$$\frac{\chi^*(G_x)}{\omega(G_x)} \leq t \qquad \text{for any integral } x \geq 0.$$

Since x is integral we can rewrite this as

$$\frac{\chi^*(G,x)}{\omega(G,x)} \leq t \qquad \text{for any integral } x \geq 0.$$

By the scaling property, we can replace "integral" by "rational" and since both χ^* and ω are continuous in x we can replace those by "real".

And by the scaling property, we can assume $\omega(G,x) \leq 1$. Summarizing these facts, we get that the conditions above are equivalent to the following:

$$\chi^*(G,x) \leq t \qquad \text{for any real } x \geq 0 \text{ for which } \omega(G,x) \leq 1.$$

Now apply Claim (1) to get the desired equivalence.

Claim (3) immediately implies the first equality of the theorem. Using Claim (2) and Claim (3) together, we get that $\mathrm{imp}(G) \leq t$ if and only if $x \in t\,\mathrm{STAB}(G)$ for any $x \in \mathrm{QSTAB}(G)$, in other words $\mathrm{QSTAB}(G) \subseteq t\,\mathrm{STAB}(G)$. This proves the second equality of the theorem.

For the third equality, use the first one.

$$\mathrm{imp}(G) = \max\{\chi^*(G,x) : x \in \mathrm{QSTAB}(G)\}$$

By the duality theorem

$$\chi^*(G,x) =$$

$$\max\{x \cdot z : z \geq 0, \sum_{v \in S} z_v \geq 1 \text{ for any stable set } S\} =$$

$$\max\{x \cdot z : z \in \mathrm{QSTAB}(\overline{G})\}$$

□

Recall that $\mathrm{STAB}(G) \subseteq \mathrm{QSTAB}(G)$ and that $\mathrm{STAB}(G) = \mathrm{QSTAB}(G)$ if and only if G is perfect.

Corollary 13.2 $\mathrm{imp} \geq 1$ *with equality if and only if G is perfect.*

Corollary 13.3 $\mathrm{imp}(G) = \mathrm{imp}(\overline{G})$

Theorem 13.4 (Gerke, McDiarmid)

$$\mathrm{imp}(G) = \lim_{k \to \infty} r_k(G).$$

Proof. First we will find a lower bound for $r_k(G)$. Let

$$r'_k(G) = \max\left\{\frac{\chi^*(G)}{\omega(G)} : x \in \mathbb{Z}^{V(G)}, x \geq 0, x_{\max} = k\right\}.$$

Clearly, $r'_k(G) \leq r_k(G)$. Pick x such that $\mathrm{imp}(G) = \frac{\chi^*(G,x)}{\omega(G,x)}$. Let $k = x_{\max}$, $l \geq k$ and let $y = \left\lfloor \frac{l}{k} x \right\rfloor$. Then $y_{\max} = l$ and so

$$r_l(G) \geq r'_l(G) \geq \frac{\chi^*(G,y)}{\omega(G,y)}.$$

Now we will prove the following inequality.

$$\chi(G_x) \leq \chi^*(G,x) + |V(G)|$$

To see this, let y be a basic optimal solution of the linear programming problem defining χ^*. Since the number of constraints is $|V(G)|$, the number of nonzero coordinates of y is at most $|V(G)|$. Then

$$\chi(G_x) \leq \sum_{S \text{ stable set}} \lceil y_S \rceil \leq \sum_S y_S + \sum_{S : y_s \neq 0} 1 \leq \chi^*(G,x) + |V(G)|.$$

Now we will find an upper bound for $r_k(G)$.

$$r_k(G) = \frac{\chi(G_x)}{\omega(G_x)} \leq \frac{\chi^*(G,x)}{\omega(G,x)} + \frac{|V(G)|}{\omega(G,x)} \leq$$

$$\frac{\chi^*(G,x)}{\omega(G,x)} + \frac{|V(G)|}{k} \leq \mathrm{imp}(G) + \frac{|V(G)|}{k}$$

Summarizing the lower and upper bounds we have

$$\frac{\chi^*(G,y)}{\omega(G,y)} \leq r_l(G) \leq \mathrm{imp}(G) + \frac{|V(G)|}{l}.$$

As $l \to \infty$, both sides tend to $\mathrm{imp}(G)$ which finishes the proof. □

Conjecture 13.5 *Every subgraph of the hexagonal lattice has imperfection ratio* 9/8*.*

Note that this is the best possible because of C_9.

14 Entropy

Definition 14.1 Let X be a random variable with distribution $p = (p_1, \ldots, p_n)$. The *entropy* of X is defined as

$$H(X) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i}.$$

Consider the following real life problem: there are two people, A and B, who want to communicate over an expensive channel. A has an information, which is a sequence of outcomes of the random variable X. A wants to send the information to B over the channel, which can carry a sequence of symbols of an alphabet. A uses a coding function (or code) to encode the message to a sequence of symbols of the alphabet. B must be able to decode the message so that B does not know where an outcome ends and where the next starts. The goal is to find a code that provides the smallest encoded messages on average. Here are the precise definitions.

Definition 14.2 Consider a random variable X and an alphabet on D symbols: $\{0, 1, \ldots, D-1\}$. A *code* C for X is a mapping

$$\text{range}(X) \to \{0, 1, \ldots, D-1\}^*.$$

The *length* of the code is

$$\sum_{x} P(x) \cdot \text{length}(C(x)).$$

A code is a *prefix code* (or *instantaneous code*) if no word is a prefix of another. Denote by $L(X)$ the length of a shortest prefix code for X.

We will see that there is a close relationship between the entropy and the shortest prefix code.

In the following examples, the alphabet is $\{0, 1\}$.

Examples

1. X is uniform over 32 elements from 0 to 31. Then $H(X) = 5$. Let C map each number to its 5-digit binary represantation. C is a prefix code, and the length of C is 5.

2. There are 8 horses on a race, the probabilities of winning are $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}\right)$. Then $H(X) = 2$. Let C map the horses to the following binary sequences respectively: (0, 10, 110, 1110, 111100, 111101, 111110, 111111). C is a prefix code and the length of C is 2.

Theorem 14.3 (Kraft inequality) *For any prefix code over an alphabet of size D with codewords of lengths $l_1, \ldots, l_m$, we have*

$$\sum_{i=1}^{m} D^{-l_i} \leq 1.$$

Conversely if $l_1, \ldots, l_m$ satisfy the inequality above, then there exists a prefix code with codewords of lengths $l_1, \ldots, l_m$.

Proof. Draw a rooted D-nary tree and label the vertices by sequences of the numbers $0, \ldots, D-1$ as the following. Label the root by the empty sequence. If a vertex is labeled by the sequence S, then the children of the vertex are labeled by $S0, S1, \ldots, S(D-1)$.

Let the depth of the tree be $l = \max\{l_1, \ldots, l_m\}$. So every codeword is among the labels, and a codeword of length l_i has D^{l-l_i} leaves as descendant. Since codewords cannot be descendant of each other and the number of leaves is D^l,

$$\sum_{i=1}^{m} D^{l-l_i} \leq D^l,$$

which proves the first part. We leave the proof of the second part as an exercise. □

Theorem 14.4

$$H(X) \leq L(X) < H(X) + 1$$

Proof. Let C be a prefix code with codewords of length $l_1, \ldots, l_m$. Let $r_i = 2^{-l_i} / \sum_{j=1}^{m} 2^{-l_j}$. Then

$$L(X) - H(X) = \sum_{i=1}^{m} p_i l_i - \sum_{i=1}^{m} p_i \log \frac{1}{p_i} =$$

$$\sum_{i=1}^{m} p_i \log \frac{p_i}{r_i} - \sum_{i=1}^{m} p_i \log \sum_{j=1}^{m} 2^{-l_j}.$$

The second term is nonpositive by the Kraft inequality, hence

$$L(X) - H(X) \geq -\sum_{i=1}^{m} p_i \log \frac{r_i}{p_i} \geq -\log \sum_{i=1}^{m} p_i \frac{r_i}{p_i} \geq 0.$$

(The second inequality follows from the convenxity of the $-\log$ function.) This proves the first part of the theorem.

For the second part, let $l_i = \lceil \log \frac{1}{p_i} \rceil$. Notice that $2^{-l_i} \leq 2^{-\log \frac{1}{p_i}}$. Then

$$\sum_{i=1}^{m} 2^{-l_i} \leq \sum_{i=1}^{m} 2^{-\log \frac{1}{p_i}} = \sum_{i=1}^{m} p_i = 1.$$

By the second part of Theorem 14.3, there exists a code with codewords of length $l_1, \ldots, l_m$. So

$$L(X) \leq \sum_{i=1}^{m} p_i l_i < \sum_{i=1}^{m} p_i \left(\log \frac{1}{p_i} + 1 \right) = H(X) + 1.$$

$\square$