RAIRO. INFORMATIQUE THÉORIQUE

PHAN DINH DIÊU

Asymptotical estimation of some characteristics of finite graphs

RAIRO. Informatique théorique, tome 11, nº 2 (1977), p. 157-174
<http://www.numdam.org/item?id=ITA_1977__11_2_157_0>

R.A.I.R.O. Informatique théorique/Theoretical Computer Science
(vol. 11, n° 2, 1977, p. 157 à 174).

ASYMPTOTICAL ESTIMATION OF SOME CHARACTERISTICS OF FINITE GRAPHS

by PHAN ĐINH DIÊU [1]

Abstract. — *Are given the asymptotical estimations of some characteristics of finite graphs. It is shown that, when* $n \to \infty$, *for almost of all graphs of n nodes every inside stable set contains at most* $[2 \log n + 1]$ *nodes, every outside stable set contains at least* $n - [2 \log n + 1]$ *nodes, every minimal cycle passes at most* $[2 \log n + 1]$ *nodes, and the chromatic number is greater than* $n/[2 \log n + 1]$. *It is proved also that when* $n \to \infty$ *for almost of all graphs of n nodes the number of inside stable sets (cliques, outside stable sets, minimal cycles) is of degree* $n^{\frac{1}{2}\log n}$.

1. INTRODUCTION

As has been well-known, a large class of combinatorial problems, in particular in the graph theory as well as the problems of finding a maximal inside stable set or a minimal outside stable set, or determining the chromatic number of a graph, etc., was known to be solvable only by very complicated algorithms. In the point of view of computational complexity theory, the presently known algorithms solving these problems require often the exponentional time, i. e., the computational time grows as an exponentional function of the length of the input. In Karp's work [2] and further papers of other authors a lot of problems in graph theory was shown to be solvable in polynomial time by nondeterministic Turing machines (i. e., to belong to the class *NP*), but if any of them is solvable in polynomial time by a deterministic machine, then so do all problems in the class *NP*. The problems of deciding whether a graph has an inside stable set of a given cardinality, or whether the chromatic number of a given graph is less than a given number, etc. belong to this class *NP*. The *NP*-completeness of these problems strengthens the conjecture that they are not solvable in polynomial time by deterministic machines.

In the relation with this problem of the computational complexity theory in this paper we shall consider the asymptotical estimation of some characteristics of graphs, that is, the inside and outside stable numbers, the chromatic number, and also the numbers of inside and outside stable sets of a graph, etc.

Some necessary definitions and notations are given in the section 2. In the section 3 it will be shown that when $n \to \infty$ for almost of all graphs of n nodes, every inside stable set contains at most $[2 \log n + 1]$ nodes, every outside

[1] Institute of Mathematics. Hanoï. Vietnam and IRIA. France.

distinct graphs with the node set X. they are denoted by $G_1, G_2, \ldots, G_p$. The k-element subsets of X are denoted by $E_1, E_2, \ldots, E_q$.

LEMMA 1 : $M\xi_{n,k} = C_n^k \,.\, 2^{-C_k^2}$.

Proof : For every graph $G_i (i = 1, \ldots, p)$ we denote by $I_k(G_i)$ the number of k-node inside stable sets of G_i, and for any set $E_j (j = 1, \ldots, q)$ we denode by $d_k(E_j)$ the number of such graphs G_i that E_j is a k-node inside stable set of G_i. It is obvious that

$$\sum_{i=1}^{p} I_k(G_i) = \sum_{j=1}^{q} d_k(E_j).$$

We now calculate the right side of this equation. Let given a subset E_j. If a graph G_i has E_j as an inside stable subset, then in G_i there are no arcs joining two nodes of E_j. Let W denote the set of all possible arcs joining two nodes in X excluding the arcs joining two nodes of E_j. It is easy to see that the cardinality of W is $|W| = C_n^2 - C_k^2$. Each graph G_i having E_j as an inside stable set corresponds to some subset of W. Therefore the number of such graphs G_i that E_j is an inside stable set is equal to

$$d_k(E_j) = 2^{C_n^2 - C_k^2}.$$

Hence,

$$\sum_{j=1}^{q} d_k(E_j) = q \,.\, 2^{C_n^2 - C_k^2} = C_n^k \,.\, 2^{C_n^2 - C_k^2}.$$

Therefore we have :

$$M\xi_{n,k} = \bar{I}_k(n) = \frac{1}{p} \sum_{i=1}^{p} I_k(G_i) = C_n^k \,.\, 2^{-C_k^2}.$$

Q.E.D.

LEMMA 2 : *Let π be a function from the graphs of n nodes into natural numbers, and let $\bar{\pi}(n)$ be the mean value of $\bar{\pi}$, i. e. :*

$$\bar{\pi}(n) = \frac{1}{p} \sum_{i=1}^{p} \pi(G_i).$$

Then the fraction of graphs G_i for which $\pi(G_i) > \bar{\pi}(n).\, z$ is less than, $1/z$.

Proof : Let A denote the number of graphs G_i such that

$$\pi(G_i) > \bar{\pi}(n) \,.\, z. \qquad (1)$$

Then we have

$$\bar{\pi}(n) = \frac{1}{p} \Sigma \pi(G_i) = \frac{1}{p} \left(\sum_{\pi(G_i) > \bar{\pi}(n).z} \pi(G_i) + \sum_{\pi(G_i) \geqslant \bar{\pi}(n).z} \bar{\pi}(G_i) \right) > \frac{1}{p} A \bar{\pi}(n) z.$$

Hence, $A/p < 1/z$, i. e., the fraction of graphs having the property (1) is less than $1/z$. Q.E.D.

From the lemmas 1 and 3 we obtain the following theorem, which gives us an asymptotical estimation of the inside stable number of finite graphs :

THEOREM 1 : *Let* $h = [2 \log n + 1]$. *When* $n \to \infty$ *for almost of all graphs* G *of* n *nodes, every inside stable set of* G *contains at most* h *nodes, therefore,* $\alpha(G) \leqslant h$.

Proof : By the lemma 1 we have $\bar{I}_{h+1}(n) = C_n^{h+1} . 2^{-C_{h+1}^2}$. It is obvious that $1/h \to 0$ as $n \to \infty$. From the lemma 3 we deduce that when $n \to \infty$ for almost of all graphs G of n nodes (or with the node set $X = \{ x_1, x_2, \ldots, x_n \}$) we have

$$I_{h+1}(G) \leqslant \bar{I}_{h+1}(n) . h,$$

i. e.,

$$I_{h+1}(G) \leqslant C_n^{h+1} . 2^{-C_{h+1}^2} . h \leqslant \frac{h . n^{h+1}}{(h + 1)! n^{h+1}} \leqslant \frac{1}{h!} \to 0.$$

This means that when $n \to \infty$ for almost of all graphs G of n nodes there are no inside stable sets of $h + 1$ nodes. Since every inside stable set of more $h + 1$ nodes must contain $(h + 1)$-node inside stable sets, we can conclude also that as $n \to \infty$ there are no inside stable sets of $h + 1$ or more nodes for almost of all graphs of n nodes. Q.E.D.

COROLLARY 1 : *Let* $h = [2 \log n + 1]$. *When* $n \to \infty$ *for almost of all graphs* G *of* n *nodes, every outside stable set of* G *contains at least* $n - h$ *nodes, therefore,* $\beta(G) \geqslant n - h$.

Proof : For each graph G_i, if E_j is an inside stable set of G_i then $\bar{E}_j = X - E_j$ is an outside stable set of G_i. Hence, for any graph G_i the number of $(n - k)$-node outside stable sets is equal to that of k-node inside stable sets. Thereby from the theorem 1 the corollary follows immediatly.

COROLLARY 2 : *When* $n \to \infty$ *for almost of all graphs* G *of* n *nodes, the chromatic number* $\gamma(G)$ *satisfies the estimation :*

$$\gamma(G) \geqslant n/[2 \log n + 1].$$

Proof : This follows from the theorem 1 and the following inequality (see [1]) :

$$\alpha(G) . \gamma(G) \geqslant n.$$

COROLLARY 3 : *When* $n \to \infty$ *for almost of all graphs* G *of* n *nodes, every clique of* G *contains at most* $h = [2 \log n + 1]$ *nodes.*

Proof : Let $X = \{ x_1, \ldots, x_n \}$. We consider the set $\mathscr{E}$ of all ordered pairs (E_1, E_2) of k-element subsets of X. For each graph $G_i (i = 1, 2, \ldots, p = 2^{C_n^2})$ with the node set X we denote by $e_k(G_i)$ the number of such pairs (E_1, E_2) that both E_1 and E_2 are k-node inside stable sets of G_i. On the other hand, for each pair $(E_1, E_2) \in \mathscr{E}$ we denote by $f_k(E_1, E_2)$ the number of such graphs G_i that E_1 and E_2 are inside stable sets of G_i. It is obvious that

$$\sum_{i=1}^{p} e_k(G_i) = \sum_{(E_1, E_2) \in \mathscr{E}} f_k(E_1, E_2). \tag{1}$$

Note that if the graph G_i has m k-node inside stable sets then $e_k(G_i) = m^2$, and the number of such graphs G_i is equal to $P_{n,k}(m)$. Therefore we have

$$\sum_{i=1}^{p} e_k(G_i) = \sum_{m=0}^{q} m^2 P_{n,k}(m).$$

where $q = C_n^k$.

Hence we obtain

$$M\xi_{n,k}^2 = \sum_{m=0}^{q} m^2 \frac{P_{n,k}(m)}{2^{C_n^2}} = \frac{1}{2^{C_n^2}} \sum_{i=1}^{p} e_k(G_i). \tag{2}$$

We now calculate the right side of (1). Let (E_1, E_2) be a pair of k-element subsets of X such that $|E_1 \cap E_2| = j$. We denote by W the set of all possible arcs joining two any nodes in X excluding the arcs joining two nodes of E_1 or two nodes of E_2. There are totally $C_n^2 - 2C_k^2 + C_j^2$ such arcs. Each graph G_i having E_1 and E_2 as two its inside stable sets corresponds to some subset of W. Therefore we have

$$f_k(E_1, E_2) = 2^{C_n^2 - 2C_k^2 + C_j^2}.$$

(note that $C_0^2 = C_1^2 = 0$). For each $j (0 \leqslant j \leqslant k)$ there are $C_n^j C_{n-j}^{k-j} C_{n-k}^{k-j}$ pairs $(E_1, E_2) \in \mathscr{E}$ such that $|E_1 \cap E_2| = j$. Note that the cardinality of $\mathscr{E}$ is $q^2 = (C_n^k)^2$. Therefore we have

$$\begin{aligned}\sum_{(E_1, E_2) \in \mathscr{E}} f_k(E_1, E_2) &= \left[(C_n^k)^2 - \sum_{j=2}^{k} C_n^j C_{n-j}^{k-j} C_{n-k}^{k-j} \right] 2^{C_n^2 - 2C_k^2} \\ &\quad + \sum_{j=2}^{k} C_n^j C_{n-j}^{k-j} C_{n-k}^{k-j} \cdot 2^{C_n^2 - 2C_k^2 + C_j^2} \\ &= \sum_{j=2}^{k} C_n^k C_k^j C_{n-k}^{k-j} 2^{C_n^2 - 2C_k^2} (2^{C_j^2} - 1) + (C_n^k)^2 2^{C_n^2 - 2C_k^2},\end{aligned} \tag{3}$$

(where we have used the equality $C_n^j C_{n-j}^{k-j} = C_n^k C_k^j$).

By virtue of (1), (2) and (3) we obtain

$$M\xi_{n,k}^2 = \frac{C_n^k}{2^{2C_k^2}} \sum_{j=2}^{k} C_k^j C_{n-k}^{k-j} (2^{C_j^2} - 1) + (M\xi_{n,k})^2.$$

Therefore,

$$D\xi_{n,k} = M\xi_{n,k}^2 - (M\xi_{n,k})^2 = \frac{C_n^k}{2^{2C_k^2}} \sum_{j=2}^{k} C_k^j C_{n-k}^{k-j} (2^{C_j^2} - 1).$$

Q.E.D.

LEMMA 5 : *Let n be a natural number such that $n \geqslant 12 \log n$. For any $k \leqslant \log n$ we have*

$$D\xi_{n,k} < \frac{k^5}{n^2} (M\xi_{n,k})^2.$$

Proof : In order to estimate $D\xi_{n,k}$ we put for every j $(2 \leqslant j \leqslant k)$:

$$a_j = C_k^j C_{n-k}^{k-j} (2^{C_j^2} - 1).$$

For any $j < k$ we have

$$\frac{a_{j+1}}{a_j} = \frac{(k-j)^2}{(j+1)(n-2k+j+1)} \cdot \frac{2^{C_{j+1}^2} - 1}{2^{C_j^2} - 1}.$$

If $j \geqslant 2$ we have

$$\frac{2^{C_{j+1}^2} - 1}{2^{C_j^2} - 1} \leqslant \frac{2^j}{(1 - 2^{C_j^2})} \leqslant 2^{j+1}.$$

On the other hand, if $j \leqslant k$ we have

$$4(k-j)^2 \leqslant 5 \,.\, 2^{k-j}.$$

Hence if $2 \leqslant j < k$ we have

$$\frac{a_{j+1}}{a_j} \leqslant \frac{2(k-j)^2 \,.\, 2^j}{3(n-2k)} = \frac{5}{6} \cdot \frac{4(k-j)^2 \,.\, 2^j}{5(n-2k)} \leqslant \frac{5}{6} \cdot \frac{2^k}{n-2k}.$$

Therefore, it follows that when $k \leqslant \log n$ for any $j (2 \leqslant j \quad k)$ we have

$$\frac{a_{j+1}}{a_j} \leqslant 1 \quad , \quad \text{i.e.,} \qquad a_{j+1} \leqslant a_j.$$

(since $n \geqslant 12 \log n$ we have $5n \leqslant 6(n - 2 \log n)$). Thus,

$$\max_{2 \leqslant j \leqslant k} a_j = a_2 = C_k^2 C_{n-k}^{k-2} (2^{C_2^2} - 1) = C_k^2 C_{n-k}^{k-2}.$$

and we obtain

$$D\xi_{n,k} \leqslant \frac{C_n^k}{2^{2C_k^2}} k C_k^2 C_{n-k}^{k-2} = \frac{k C_k^2 C_{n-k}^{k-2}}{C_n^k} (M\xi_{n,k})^2.$$

We have

$$\frac{k C_k^2 C_{n-k}^{k-2}}{C_n^k} = \frac{k \cdot k(k-1)(n-k) \dots (n-2k+3) \cdot k!}{2n(n-1) \dots (n-k+1) \cdot (k-2)!} < \frac{k^3(k-1)^2}{2n(n-1)} < \frac{k^5}{n^2}.$$

Therefore we obtain finally :

$$D\xi_{n,k} < \frac{k^5}{n^2} (M\xi_{n,k})^2.$$

Q.E.D.

LEMMA 6 : *Let* $k = [\log n]$. *When* $n \to \infty$ *for almost of all graph* G *of* n *nodes, the number* $I_k(G)$ *of* k*-node inside stable sets satisfies the following estimation :*

$$\frac{C_n^k}{2^{C_k^2}}\left(1 - \frac{k^3}{n}\right) < I_k(G) < \frac{C_n^k}{2^{C_k^2}}\left(1 + \frac{k^3}{n}\right).$$

Proof : By using the Tchebychev's inequality in the probability theory we have for any $t > 0$:

$$P(|\xi_{n,k} - M\xi_{n,k}| \geqslant t) < \frac{D\xi_{n,k}}{t^2},$$

where $P(A)$ is the probability of the event A. Taking $t = \frac{k^3}{n} M\xi_{n,k} = \frac{k^3}{n} \frac{C_n^k}{2^{C_k^2}}$ we have

$$\frac{D\xi_{n,k}}{t^2} < \frac{k^5 (M\xi_{n,k})^2}{n^2} \cdot \frac{n^2}{k^6 (M\xi_{n,k})^2} = \frac{1}{k} \underset{n \to \infty}{\to} 0.$$

Therefore, if $k = [\log n]$ then as $n \to \infty$ for almost of all graph G with the node set X we have

$$\left| I_k(G) - \frac{C_n^k}{2^{C_k^2}} \right| < \frac{k^3}{n} \cdot \frac{C_n^k}{2^{C_k^2}}.$$

($\xi_{n,k}$ is the stochastic variable taking the value $I_k(G)$ for each given graph G with the same probability $2^{-C_n^2}$).

Q.E.D.

THEOREM 3 : *When $n \to \infty$ for almost of all graphs G with the node set $X = \{ x_1, x_2, \ldots, x_n \}$ (i. e., of n nodes), the number $I(G)$ of all inside stable sets satisfies the estimation :*

$$n^{\frac{1}{2}\log n - \log\log n - 1} < I(G) < n^{\frac{1}{2}\log n + \log\log n + 2}.$$

Proof : 1) Put $k = [\log n]$. When n is enough large we have $C_n^k \geqslant (n/k)^k$, because $(n - k + i)/i > n/k$ for $i \leqslant k$. Hence

$$C_n^k \geqslant \left(\frac{n}{k}\right)^k \geqslant \left(\frac{n}{\log n}\right)^{\log n - 1} \geqslant n^{\log n - \log\log n - 1},$$

$$2^{C_k^2} \leqslant (2^{\log n - 1})^{\frac{1}{2}\log n} = \left(\frac{n}{2}\right)^{\frac{1}{2}\log n}.$$

Therefore, from the lemma 6 we deduce that as $n \to \infty$ for almost of all graphs G of n nodes we have

$$I_k(G) > \frac{C_n^k}{2^{C_k^2}}\left(1 - \frac{k^3}{n}\right) \geqslant \frac{n^{\log n - \log\log n - 1} n^{\frac{1}{2}}}{n^{\frac{1}{2}\log n}}\left(1 - \frac{\log^3 n}{n}\right)$$
$$> n^{\frac{1}{2}\log n - \log\log n - 1}.$$

Obviously that $I(G) > I_k(G)$, and thus the first inequality of the theorem is proved.

2) In order to find an upper bound of $I(G)$ for almost of all graphs G of n nodes we may use the second inequality in the lemma 6. But it is better to estimate immediatly the values of $\bar{I}_k(n)$. By the lemma 1 we have

$$\bar{I}_k(n) = C_n^k \cdot 2^{-C_k^2}.$$

Therefore, for any $k < n$:

$$\frac{\bar{I}_{k+1}(n)}{\bar{I}_k(n)} = \frac{n - k}{2^k(k + 1)}$$

If $k \leqslant \log n - 1$ we have

$$\frac{\bar{I}_{k+1}(n)}{\bar{I}_k(n)} \geqslant \frac{2(n - \log n + 1)}{n \cdot \log n} > \frac{1}{\log n}$$

If $k > \log n - 1$ we have

$$\frac{\bar{I}_{k+1}(n)}{\bar{I}_k(n)} < \frac{2(n - \log n + 1)}{n \cdot \log n} < \frac{2}{\log n}.$$

By putting $h = [\log n]$ we have

$$\bar{I}_k(n) < \begin{cases} (\log n)^{h-k}\bar{I}_h(n) & , \text{ if } k < h, \\ \left(\dfrac{2}{\log n}\right)^{k-h} \bar{I}_h(n) & , \text{ if } k \geqslant h. \end{cases}$$

Let us denote $\bar{I}(n)$ the mean value of the number of inside stable sets $I(G)$ taken for all graphs G of n nodes. Then, when n is enough large we have

$$\begin{aligned} \bar{I}(n) &= \frac{1}{2^{C_n^2}} \sum_{i=1}^{p} \bar{I}(G_i) = \sum_{k=1}^{n} \bar{I}_k(n) \\ &< \sum_{k<h} \bar{I}_h(n) \, . \, (\log n)^{h-k} + \sum_{k \geqslant h} \bar{I}_h(n) \, . \left(\frac{2}{\log n}\right)^{h-k} \\ &\leqslant \bar{I}_h(n)\left(\log n^{\log n} + \frac{1}{1\text{-}2/\log n}\right) \\ &\leqslant \bar{I}_h(n) \, . \, (n^{\log \log n} + 2). \end{aligned}$$

On the other hand, since $\log n - 1 < h \leqslant \log n$ we have

$$C_n^h \leqslant \frac{n^h}{h!} \leqslant \frac{n^{\log n}}{2^{\log n}} = n^{\log n - 1},$$

$$C_h^2 > \frac{(\log n - 1)(\log n - 2)}{2} > \log n\left(\frac{1}{2}\log n - 2\right),$$

$$2^{C_h^2} > n^{\frac{1}{2}\log n - 2}$$

Hence,

$$\bar{I}_h(n) = C_n^h \, . \, 2^{-C_h^2} < n^{\frac{1}{2}\log n + 1}$$

Therefore,

$$\begin{aligned} \bar{I}(n) &< n^{\frac{1}{2}\log n + 1}(n^{\log \log n} + 2) \\ &< 2n^{\frac{1}{2}\log n + \log \log n + 1} \end{aligned}$$

By the lemma 2 the fraction of graphs G of n nodes for which

$$I(G) > \bar{I}(n) \, . \, n/2$$

is less than $2/n$. Thus, when $n \to \infty$ for almost of all graphs G of n nodes we have

$$I(G) < \bar{I}(n) \cdot \frac{n}{2} < n^{\frac{1}{2}\log n + \log\log n + 2}.$$

Q.E.D.

Let $G = (X, E)$ be a graph. An inside stable set $B \subseteq X$ is called maximal, if it is not contained in any other inside stable set of G. Let us denote $IM(G)$ the number of maximal inside stable sets of G, and $IM_k(G)$ the number of k-node maximal inside stable sets of G. We have the following asymptotical estimation for $IM(G)$:

THEOREM 4 : *When $n \to \infty$ for almost of all graphs G of n nodes, the number of maximal inside stable sets satisfies the estimation :*

$$n^{\frac{1}{2}\log n - \log\log n - 3} < IM(G) < n^{\frac{1}{2}\log n + \log\log n + 2}.$$

Proof : Since the number of maximal inside stable sets of a graph is not greater than that of inside stable sets of it, the second inequality for the upper bound follows immediatly from the theorem 3. In order to estimate the lower bound we put $k = [\log n]$ and $h = [2\log n + 1]$. By the theorem 1 as $n \to \infty$ every inside stable set of G does not contain more than h nodes for almost of all graphs G of n nodes. Every maximal inside stable set having not more h nodes contains at most C_h^k k-node inside stable sets, and on the other hand, every k-node inside stable set must be contained in some maximal inside stable set. Therefore, as $n \to \infty$ for almost of all graphs G of n nodes, the number of maximal inside stable sets is not less than $1/C_h^k$ times of the number of k-node inside stable sets, i. e.,

$$IM(G) \geqslant \frac{1}{C_h^k} I_k(G).$$

As in the proof of the theorem 3, as $n \to \infty$ for almost of all graphs G of n nodes we have

$$I_k(G) > n^{\frac{1}{2}\log n - \log\log n - 1}.$$

On the other hand,

$$C_h^k \leqslant 2^{h-1} \leqslant 2^{2\log n} = n^2.$$

Thus, as $n \to \infty$ for almost of all graphs G of n nodes we have

$$\overline{IM}(G) > n^{\frac{1}{2}\log n - \log\log n - 3}.$$

Q.E.D.

REMARK : We can find an asymptotical estimation for $IM(G)$ by starting from the estimation of $IM_k(n)$, the mean value of the number of k-node maximal inside stable sets of the graphs of n nodes. By a computation as well as in the case of $I_k(n)$, we can obtain :

$$\overline{IM}_k(n) = \frac{1}{2^{C_n^2}} \cdot C_n^k \cdot 2^{C_n^2 - k}(2^k - 1)^{n-k}.$$

But by the method used above, this formula does not give us the better estimation than that obtained in the theorem 4.

We note that for any graph G, the number of (minimal) outside stable sets is equal to that of (maximal) inside stable sets, and the number of (maximal) cliques of a graph $G = (X, E)$ is equal to that of (maximal) inside stable sets of the graph $\bar{G} = (X, \bar{E})$. Therefore, from the theorems 3 and 4 we obtain the following corollaries :

COROLLARY 5 : *When $n \to \infty$ for almost of all graphs G of n nodes, the number $O(G)$ of outside stable sets and the number $Om(G)$ of minimal outside stable sets satisfy the estimations :*

$$n^{\frac{1}{2}\log n - \log\log n - 1} < O(G) < n^{\frac{1}{2}\log n + \log\log n + 2},$$

$$n^{\frac{1}{2}\log n - \log\log n - 3} < Om(G) < n^{\frac{1}{2}\log n + \log\log n + 2}.$$

COROLLARY 6 : *When $n \to \infty$ for almost of all graphs G of n nodes, the number $K(G)$ of cliques and the number $KM(G)$ of maximal cliques satisfy the estimations :*

$$n^{\frac{1}{2}\log n - \log\log n - 1} < K(G) < n^{\frac{1}{2}\log n + \log\log n + 2},$$

$$n^{\frac{1}{2}\log n - \log\log n - 3} < KM(G) < n^{\frac{1}{2}\log n + \log\log n + 2}.$$

6. ASYMPTOTICAL ESTIMATION OF THE NUMBER OF MINIMAL CYCLES OF A GRAPH

In order to estimate the number $C(G)$ of minimal cycles of a graph G, first of all we calculate the dispersion $D\zeta_{n,k}$ of the stochastic variable $\zeta_{n,k}$ which takes the value m with the probability $r_{n,k}(m)$, i. e., the probability of the fact that a graph of n nodes has m k-node minimal cycles.

LEMMA 7 : $D\zeta_{n,k} \leqslant (M\zeta_{n,k})^2 \cdot \dfrac{1}{C_n^k} \sum_{j=2}^{k} C_k^j C_{n-k}^{k-j}(2^{C_j^2} - 1).$

Proof : Let $X = \{x_1, \ldots, x_n\}$. As well as in the proof of the lemma 4 we consider the set $\mathscr{E}$ of all ordered pairs (E_1, E_2) of k-element subsets of X. For each graph G_i $(i = 1, 2, \ldots, p)$ with the node set X we denote by $e_k(G_i)$ the number of such pairs (E_1, E_2) that G_i has two minimal cycles passing the nodes of E_1 and E_2 respectively. And for each pair $(E_1, E_2) \in \mathscr{E}$ we denote by $f_k(E_1, E_2)$ the number of such graphs G_i that E_1 and E_2 are two its minimal cycles. It is obvious that

$$\sum_{i=1}^{p} e_k(G_i) = \sum_{(E_1, E_2) \in \mathcal{E}} f_k(E_1, E_2).$$

Similarly as has been proved in the lemma 4 for the case of $M\zeta_{n,k}^2$ we can show also that

$$M\zeta_{n,k}^2 = \frac{1}{2^{C_n^2}} \Sigma e_k(G_i) = \frac{1}{2^{C_n^2}} \sum_{(E_1, E_2) \in \mathcal{E}} f_k(E_1, E_2).$$

Let $(E_1, E_2) \in \mathscr{E}$ and $|E_1 \cap E_2| = j$. We denote by W the set of all possible arcs joining two any nodes in X excluding the arcs joining two nodes of E_1 or two nodes of E_2. We have $|W| = C_n^2 - 2C_k^2 + C_j^2$. Let $G_i = (X, A_i)$ be a graph with the node set X such that E_1 and E_2 are two its minimal cycles. The set of arcs A_i can be divided into $A_i = A_{i_1} \cup A_{i_2}$, where $A_{i_1} \subseteq W$ and A_{i_2} is the set of arcs in G_i joining two nodes of E_1 or of E_2. It is obvious that there are totally $2^{|W|}$ possible choices for A_{i_1}. As from E_1 (or E_2) one can form $k!/2k$ distinct k-node cycles, there are at most $(k!/2k)^2$ possible choices for A_{i_2} (there are exactly $(k!/2k)^2$ possible choices for A_{i_2} when $|E_1 \cup E_2| \leqslant 1$). Thus we have

$$f_k(E_1, E_2) \leqslant \left(\frac{k!}{2k}\right)^2 . 2^{C_n^2 - 2C_k^2 + C_j^2}.$$

Note that $|\mathscr{E}| = q^2 = (C_n^k)^2$, and for each j $(0 \leqslant j \leqslant k)$ there are $C_n^j C_{n-j}^{k-j} C_{n-k}^{k-j}$ pairs $(E_1, E_2) \in \mathscr{E}$ such that $|E_1 \cap E_2| = j$. Therefore we obtain

$$\begin{aligned} M\zeta_{n,k}^2 &= \frac{1}{2^{C_n^2}} \Sigma f_k(E_1, E_2) \\ &\leqslant \left[(C_n^k)^2 - \sum_{j=2}^{k} C_n^j C_{n-j}^{k-j} C_{n-k}^{k-j}\right]\left(\frac{k!}{2k}\right)^2 . 2^{-2C_k^2} \\ &+ \sum_{j=2}^{k} C_n^j C_{n-j}^{k-j} C_{n-k}^{k-j} \left(\frac{k!}{2k}\right)^2 . 2^{-2C_k^2 + C_j^2} \\ &\leqslant (M\zeta_{n,k})^2 + \left(\frac{k!}{2k}\right)^2 \sum_{j=2}^{k} C_n^j C_{n-j}^{k-j} C_{n-k}^{k-j} . 2^{-2C_k^2}(2^{C_j^2} - 1) \\ &\leqslant (M\zeta_{n,k})^2 + \left(\frac{k!}{2k}\right)^2 \frac{C_n^k}{2^{2C_k^2}} \sum_{j=2}^{k} C_k^j C_{n-k}^{k-j}(2^{C_j^2} - 1). \end{aligned}$$

Hence,

$$D\zeta_{n,k} \leqslant (M\zeta_{n,k})^2 \cdot \frac{1}{C_n^k} \sum_{j=2}^{k} C_k^j C_{n-k}^{k-j} (2^{C_j^2} - 1).$$

Q.E.D.

By the same reasoning used in the proofs of the lemmas 5 and 6 we can prove that when n is enough large, for any $k \leqslant \log n$ we have

$$D\zeta_{n,k} < \frac{k^5}{n^2} (M\zeta_{n,k})^2,$$

and also the following lemma :

LEMMA 8 : *Let $k = [\log n]$. When $n \to \infty$ for almost of all graphs G of n nodes, the number $C_k(G)$ of k-node minimal cycles satisfies the inequality :*

$$\frac{k!}{2k} \cdot \frac{C_n^k}{2^{C_k^2}} \left(1 - \frac{k^3}{n}\right) < C_k(G) < \frac{k!}{2k} \frac{C_n^k}{2^{C_k^2}} \left(1 + \frac{k^3}{n}\right).$$

THEOREM 5 : *When $n \to \infty$ for almost of all graphs G of n nodes, the number $C(G)$ of all minimal cycles satisfies the estimation :*

$$n^{\frac{1}{2}\log n - \frac{1}{2}\log\log n - 1} < C(G) < n^{\frac{1}{2}\log n + \log\log n + 2}.$$

Proof : 1) Put $k = [\log n]$. By the lemma 8 when $n \to \infty$ for almost of all graphs G of n nodes we have

$$C(G) > C_k(G) > \frac{k!}{2k} \cdot \frac{C_n^k}{2^{C_k^2}} \left(1 - \frac{k^3}{n}\right).$$

As has been shown in the proof of the theorem 3 we have

$$\frac{C_n^k}{2^{C_k^2}} \left(1 - \frac{k^3}{n}\right) > n^{\frac{1}{2}\log n - \log\log n - \frac{1}{2}} \left(1 - \frac{\log^3 n}{n}\right)$$
$$> n^{\frac{1}{2}\log n - \log\log n - 1} \cdot 2(\log n + 1).$$

On the other hand we have

$$k! > 2^{\frac{k}{2}} \cdot \left(\frac{k}{2}\right)^{\frac{k}{2}} > k^{\frac{k}{2}} > (\log n)^{\frac{1}{2}\log n} = n^{\frac{1}{2}\log\log n}.$$

Therefore, as $n \to \infty$ for almost of all graphs C of n nodes we have

$$C(G) > n^{\frac{1}{2}\log n - \frac{1}{2}\log\log n - 1}.$$

2), By the lemma 3 we have

$$\bar{C}_h(n) = C_n^h \cdot 2^{C_h^2} \frac{h!}{2h}.$$

Therefore, for any $h < n$:

$$\frac{\bar{C}_{h+1}(n)}{\bar{C}_h(n)} = \frac{h(n-h)}{2^h(h+1)}$$

If $3 \leqslant h \leqslant \log n - 1$ we have

$$\frac{\bar{C}_{h+1}(n)}{\bar{C}_h(n)} \geqslant \frac{h}{h+1} \frac{2(n - \log n + 1)}{n} > 1.$$

If $\log n \leqslant h < n$ we have

$$\frac{\bar{C}_{h+1}(n)}{\bar{C}_h(n)} \leqslant \frac{h}{h+1} \frac{(n - \log n)}{n} < 1.$$

By putting $k = [\log n]$ we have for any $h\,(3 \leqslant h \leqslant n)$:

$$\bar{C}_h(n) \leqslant \bar{C}_k(n) \leqslant C_n^k \cdot 2^{C_k^2} \cdot \frac{k!}{2k} \leqslant n^{\frac{1}{2}\log n + 1} \cdot \frac{k!}{2k}$$

$$\leqslant \frac{1}{\log n} \cdot n^{\frac{1}{2}\log n + \log\log n + 1}.$$

Let us denote $\bar{C}(n)$ the mean value of the number of minimal cycles $C(G)$ taken for all graphs G of n nodes. Then we have :

$$\bar{C}(n) = \frac{1}{2^{C_n^2}} \sum_{i=1}^{p} C(G_i) = \sum_{h=1}^{n} \bar{C}_h(n)$$

$$\leqslant \frac{1}{\log n} n^{\frac{1}{2}\log n + \log\log n + 2}.$$

By virtue of the lemma 2 we deduce that the fraction of graphs G of n nodes for which $C(G) > \bar{C}(n) . \log n$ is less than $1/\log n$. Thus, when $n \to \infty$ for almost of all graphs G of n nodes we have :

$$C(G) < \bar{C}(n) . \log n < n^{\frac{1}{2}\log n + \log\log n + 2}.$$

Q.E.D.

7. CONCLUSION

As has been proved above, the number of inside (outside) stable sets of a graph of n nodes is of degree $n^{\frac{1}{2}\log n}$ for almost of all such graphs as $n \to \infty$. Thereby it follows that every algorithm listing all inside (outside) stable sets of a graph must work at least in $n^{\frac{1}{2}\log n}$-time. The problem of deciding if a given graph has an inside stable set of a given cardinality is shown to be NP-complete (see [2]). If one has to find all inside stable sets of $k - 1$ nodes before deciding whether the graph has an inside stable set of k nodes, then every algorithm solving the mentionned above problem must work at least in $n^{\frac{1}{2}\log n}$-time, i. e., it cannot work in polynomial time. These fact leads us to the following conjectures :

1) Each NP-complete problem cannot be solvable in polynomial time by a deterministic machine. Every algorithm solving such a problem requires at least $n^{a \log n}$-time for some positive constant a (where n is the length of the input of the given problem).

2) Each NP-complete problem can be solvable in $n^{\log^b n}$-time by a deterministic machine, at least for almost of all inputs of it, where b is a suitable constant. Note that the function $n^{\log^b n}$ is quite lower than the exponentional function 2^n as $n \to \infty$.

REFERENCES

1. C. Berge, *Théorie des graphes et ses applications*, Dunod, Paris, 1958.
2. R. M. Karp, *Reducibility among combinatorial problems.* Technical Report n° 3, April 1972, Department of Computer Science, University of California, Berkeley.
3. V. V. Glagolev, *Some estimations of disjunctive normal forms of boolean functions.* Problems of Cybernetics, v. 19, Moscow, 1967, p. 75-94 (in Russian).
4. Phan Dinh Diêu, *On the complexity of the problem of state minimization for incompletely specified sequential automata.* Elektronische Informationsverarbeitung und Kybernetik, **11**, 1975, p. 3-17.