# COLORING A GRAPH BY SOCIAL INFLUENCE

## A SIMPLE MODEL OF COLOR CHANGES


**Moses A. Boudourides**

Department of Mathematics

University of Patras

265 00 Rio-Patras, Greece

mboudour@upatras.gr



**For presentation at the XXIV International Sunbelt Social Network Conference**

**Portorož, Slovenia, May 12-16, 2004**

**Work in progress – First draft of April 25, 2004**

**(Not reviewed by a native speaker)**

**Please do not quote without the authors' permission**



**Abstract** [To be written]


**Introduction** [To be written]

**Fundamental Concepts**

Let $G = (V,E)$ be an **undirected graph** with **vertex** set $V = \{1,2, \dots, N\}$ and **edge** set $E$. $G$ is supposed to be **simple** without self-loops or multiple edges and **connected** in the sense that any two vertices are connected by a chain of edges, also called **links**. A **subgraph** $H$ of $G$ is a graph $H = (U,F)$ such that $U \subseteq V$ and $F \subseteq E$ and the assignment of endpoints to edges in $H$ is the same as in $G$. Let $V(i)$ denote the set of all vertices which are **adjacent** to $i$ (connected with a single edge) and let $\deg(i) = |V(i)|$ be the **degree** of $i$ (i.e., the number of edges incident to $i$).

Suppose that the symbols 0 and 1 denote two colors and $S = \{0,1\}$. Then a **2-coloring** or just **coloring** of the graph $G$ (which is then called 'colored graph') is a labeling $C: V \to S$, by which a color is assigned to each vertex of the graph. When $C(i) = s \in S$, we say that "the color of $i$ is $s$" or that "$i$ is colored $s$." A set of vertices $U \subseteq V$ is called **monochromatic** if $C(i) = s$, for all $i \in U$, and **dichromatic** if there exist $i, j \in U$, $i \neq j$, with $C(i) \neq C(j)$. Since a link corresponds to two vertices, we can also talk about monochromatic or dichromatic links. When $G$ is colored, the vector $\varphi = (C(1), C(2), \dots, C(N)) \in S^N$ ($|S^N| = 2^N$) is called **graph coloring** (or **configuration** or **profile** of colors on the graph).

Clearly, every graph coloring creates a partition of $G$ into two subgraphs $G^0 = (V^0,E^0)$ and $G^1 = (V^1,E^1)$, where $V = V^0 \cup V^1$, $V^0 \cap V^1 = \emptyset$, $E = E^0 \cup E^1 \cup E^{01}$ and $E^0$, $E^1$, $E^{01}$ are pairwise disjoint subsets of $E$. The subgraphs $G^0$, $G^1$ are called **monochromatic components** of $G$, links $E^0$, $E^1$ in each of them are monochromatic but links $E^{01}$ between them are dichromatic. Of course, $G^0$, $G^1$ need not be connected and one of them might be possibly empty. Apparently, $|V^s| = N^s$, $s \in S$, and $\Sigma_{s \in S} N^s = N$. Furthermore, let us denote by $V^s(i)$ the set of all $s$-colored neighbors of vertex $i$ and by $n^s(i) = |V^s(i)|$ their total number. Clearly, $\bigcup_{s \in S} V^s(i) = V(i)$ and $\sum_{s \in S} n^s(i) = \deg(i)$, for any $i \in V$.

## Interpretation of Colored Graphs

Considering that the set of vertices of a graph represents a population of actors and their links represent relationships between actors, then we might interpret a graph coloring in a variety of ways. For instance, in the context of ferromagnetism and, in particular of the *Ising model* (Prum & Fort 1991), the actors might constitute a system of magnetized atoms situated at the vertices of a regular lattice and directed towards opposite orientations, usually denoted by +1 and -1. This is a stochastic model and it is considered to belong into the general category of theories of *infinite particle systems* or *interacting particle systems* (Durrett 1981; Griffeath 1979; Liggett 1985). A special case of the latter theories includes the *voter model* (Holley & Liggett 1975), in which actors might be conceived as voters possessing just one between two opinions on a fixed issue, usually symbolized by the states 0 and 1.

Unlike natural systems with non-human actors, in social sciences – as it happens within the standard setting of social networks analysis (Wasserman & Faust 1994) – actors are conceived to be either individuals or aggregate units (such as groups of individuals or organizations or any other collective identities). Then the colors are considered to be either attributes or statuses that actors possess or states which are assigned to actors. Again these models can be described as Markov chains or other such stochastic processes (Spitzer 1970; Kindermann & Snell 1980; Brémaud 1999). These models appear in studies of contact processes (Harris 1974), epidemics (Keeling 1999) and spread of infectious diseases (Bailey 1975). The corresponding propagation processes in social systems include threshold models of collective behavior (Granovetter 1978), social contagion (Morris 2000) and global cascades (Watts 2002). Another similar stochastic process appears in diffusion theories, as in the case of the adoption of one between two competing technologies (Arthur 1989).

From all the above diverse applications, it is clear that there exist multiple contexts in which the “color” of a vertex/actor can be interpreted. It can mean “decision” or “opinion” or “action” or “activation” or “choice” or “adoption” or something similar.

However, in all these cases, whatever is happening is restricted among a finite number of certain mutually exclusionary alternatives. An actor is responsible of her agencies but she might be influenced by her neighborhood too. This is the situation that we will examine in the next section through a simple model taking into account the network effect of social influence.

Finally, let us remark that, in the case of a 2-coloring, there is a natural way to assign a sign on all the links of *G*: monochromatic links are positively signed and dichromatic links are negatively signed. Clearly, the product of signs on all cycles of a dichromatic graph *G* is positive, i.e., such graph is 'balanced' in the sense of balance theory, which was initiated by Heider (1946). Then the structure theorem of balanced graphs (Davis 1967) is completely consistent with the partition of the set of vertices of a dichromatic graph into two blocks both with monochromatic (thus, positive) links in their interior and dichromatic (thus, negative) links inter-connecting them.

**A Simple Model of Color Changes Driven by Social Influence**

Now we are going to allow vertices to change their colors by taking into account the colors of their neighbors – this is what we mean by 'network influence.' However, when vertices may exhibit such an active characteristic, then one is entitled to call them actors and to perceive the graph they constitute as a social network over which they are embedded. In this sense, network influence becomes a genuine social influence.

So, let us assume that at some time every actor possesses a certain color (which does not have to be the same for everybody). It is reasonable to assume that every actor 'knows' her own color but she can also 'see' the colors of her neighbors. Moreover, every actor is free to change (or update) her color after she receives information about the distribution of colors in her neighborhood. However, her decision – as to whether she wishes to change or maintain her color – is not taken arbitrarily but it strictly follows certain rules. In fact, these rules prescribe the concrete ways that an actor might evaluate the input from her neighbors. In other words, these rules express the criteria that actors are following when they make up their minds about possible changes of their colors after they have been influenced by their neighbors' preferences.

The very idea of the 'decision criteria' that we are assuming is that first an actor compares the number of his or her neighbors who are in a certain color with his susceptibility threshold towards this color and then, accordingly, he or she decides which color to adopt. Therefore, to each actor $i$ there correspond two numbers $\theta^o(i)$, $\theta^1(i) \in (0,1)$ called **thresholds** such that $\theta^o(i) + \theta^1(i) = 1$ and these numbers represent the susceptibility of actor $i$ towards color 0 and 1 (respectively) in the following sense:

$$\text{if } n^o(i) \geq \theta^o(i)\deg(i), \text{ then actor } i \text{ adopts color 0, i.e., } C(i) = 0;$$
$$\text{or, equivalently, if } n^1(i) > \theta^1(i)\deg(i), \text{ then actor } i \text{ adopts color 1, i.e., } C(i) = 1.$$

This means that actor $i$ accepts color 0 whenever at least $\theta^o(i)\deg(i)$ of his or her neighbors are already in color 0 (or, equivalently, whenever at most $\theta^1(i)\deg(i)$ of the neighbors are already in color 1) and he or she accepts color 1 whenever his or her neighbors in color 0 are (strictly) less than $\theta^o(i)\deg(i)$ (or, equivalently, the neighbors in color 1 are (strictly) more than $\theta^1(i)\deg(i)$).

Using the fact that $n^o(i) + n^1(i) = \deg(i)$, these 'decision criteria' can be reformulated as follows:

$$\text{if } n^o(i) = 0, \text{ then } C(i) = 1;$$
$$\text{if } n^o(i) \neq 0 \text{ and } n^1(i) / n^o(i) \leq \theta^1(i) / \theta^o(i), \text{ then } C(i) = 0.$$

Or, equivalently:

$$\text{if } n^1(i) = 0, \text{ then } C(i) = 0;$$
$$\text{if } n^1(i) \neq 0 \text{ and } n^o(i) / n^1(i) \leq \theta^o(i) / \theta^1(i), \text{ then } C(i) = 1.$$

In other words, when both $n^{o}(i)$, $n^{1}(i) \neq 0$, denoting:

$$\rho^{o1}(i) = \frac{n^{1}(i)}{n^{0}(i)},$$

$$\tau^{o1}(i) = \frac{\theta^{1}(i)}{\theta^{0}(i)},$$

the 'decision criteria' become:

$$\rho^{o1}(i) \leq \tau^{o1}(i) \Rightarrow C(i) = 0$$

and, equivalently, $\rho^{o1}(i) > \tau^{o1}(i) \Rightarrow C(i) = 1$.

Furthermore, let us remark that, although, by definition, 0 and 1 are excluded from the range of values that thresholds take, this is what would happen if the thresholds could have taken these values:

- When $\theta^{o}(i) = 0$ in which case $\theta^{1}(i) = 1$, then actor $i$ would always have to be colored 0, independently of the color of the neighbors.
- However, when $\theta^{o}(i) = 1$ corresponding to $\theta^{1}(i) = 0$, then actor $i$ would have to be colored 0 only when all the neighbors were colored 0 and, if there existed a single neighbor in color 1, then actor $i$ would be colored 1.

Thus, in this model, although every actor $i$ might have her own (distinctive) thresholds $\theta^{o}(i)$, $\theta^{1}(i)$ ($\theta^{o}(i) + \theta^{1}(i) = 1$), depending on the value of her thresholds, an actor might have different proclivities towards the two colors 0 and 1. This is why an actor $i$ might be called **0-philic**, when $0 < \theta^{o}(i) < 1/2$, and **1-philic**, when $1/2 < \theta^{o}(i) < 1$ (i.e., when $0 < \theta^{1}(i) < 1/2$). The case $\theta^{o}(i) = \theta^{1}(i) = 1/2$ is rather ambivalent depending on whether the degree of $i$ is odd or even. Furthermore, we should distinguish between the cases of homogeneous and heterogeneous thresholds over the whole graph: We say that the distribution of thresholds is **homogeneous** over the graph $G$ if $\theta^{o}(i) = \theta^{o}$ (equivalently,

$\theta^1(i) = 1 - \theta^o$), for all $i \in V$, and for some $\theta^o \in (0,1)$, and **heterogeneous** otherwise (i.e., if $\theta^o(i)$, $\theta^1(i)$ vary in $i$).

### Invariance and Structure of Colorings

Now the question is up to when are actors supposed to update their colors if they do this by conforming to the rules set in previous section? Stated in an equivalent way, to answer we need to determine which the **invariant** (or **equilibrium**) graph colorings are: these are colorings with colors that never change, i.e., colorings which might be interpreted as having reached an (invariable) equilibrium.

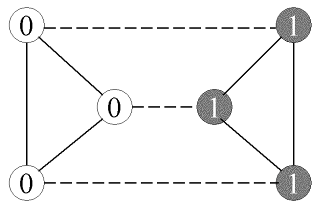
Clearly, an arbitrary graph coloring is invariant only as far as the two components in the partition $V = V^0 \bigcup V^1$ can maintain their colors according to the above criteria. This means that the monochromatic components $V^0$, $V^1$ of $G$ are invariant if and only if:

$$\text{for all } i \in V^0, \text{ either } n^1(i) = 0 \text{ or } \rho^{o1}(i) \leq \tau^{o1}(i);$$
$$\text{or, equivalently, for all } i \in V^1, \text{ either } n^o(i) = 0 \text{ or } \rho^{o1}(i) > \tau^{o1}(i).$$

### Examples of Invariant Colorings and Interpretation of Thresholds

Apparently, any homogeneous monochromatic graph coloring (i.e., all actors have either color 0 or color 1) is invariant. But there exist heterogeneous invariant graph colorings, which are dichromatic, as in the following example.

**Example 1:** Let us consider the 3-regular graph of 6 actors half of them are white and half are black. We want to determine what range of thresholds $\theta^o(i)$ renders this coloring invariant




.

As in all graph drawings, white vertices will be in color 0, black vertices in color 1, black links are monochromatic and dashed links are dichromatic.

For any white vertex $i \in V^0$, $n^o(i) = 2$ and $n^1(i) = 1$, while, for any black vertex $i \in V^1$, $n^1(i) = 2$ and $n^0(1) = 1$. Therefore, the above invariance conditions imply that $\theta(i) > 1/3$ on $V^0$, while $\theta(i) \leq 2/3$ on $V^1$. Thus, when $\theta(i) \in (1/3,2/3]$, for all $i$, the above dichromatic coloring is invariant.■

Next, we would like to explore the structure of invariant monochromatic blocks when all actors have a common threshold value $\theta(i) = \theta \in (0,1]$, for all $i$. In what follows **IMB** stands for 'invariant monochromatic block.' First we claim:

**Proposition 1.** When all thresholds are the same, any non-trivial IMB should be composed of at least three connected actors.

**Proof:** Let us consider a white block (similarly otherwise). If $|V^0| = 1$, say $V^0 = \{i\}$, then $n^0(i) = 0$ and the invariance conditions imply that $\theta > 1$, which is a contradiction. If $|V^0| = 2$, necessarily the two 0-colored vertices should be connected and $V^1$ should be composed of at least two connected vertices. Let $i \in V^0$ be a vertex with at least one dichromatic link. Then $n^0(i) = 1$, $n^1(i) = \kappa \geq 1$ and, so, $\theta > \kappa/(1+\kappa)$. But there always exist **[???]** $j \in V^1$ with $n^0(j) = 1$, $n^1(j) \leq \kappa - 1$ and, so, $\theta < (\kappa\text{-}1)/\kappa$, which is again a contradiction.■

**Proposition 2**. There is no dichromatic coloring of a $d$-lattice with even degree when all actors have equal thresholds.

**Proof:** Without any loss of generality, let us assume that the lattice is a $2k$-regular graph of $m \geq k(k + 1)$ vertices, which are arranged on a circle in such a way that each vertex is joined with the $k$ nearest vertices in each direction of the circle. By Proposition 1, it suffices to consider a triangular IMB composed of three connected vertices, which are the only ones in color (say) 0 in the graph. Then, on the middle of these three vertices $n^0 = 2$, $n^1 = 2k - 2$, which implies that $\theta > (k - 1)/k$ (by the invariance conditions). However, there exists **[???]** a vertex in color 1 such that $n^0 = 2$, $n^1 = 2k - 2$ and, so (since the latter vertex needs to preserve color 1), $\theta \leq (k - 1)/k$, which is a contradiction.■

[Interpretation of threshold should be written here.]

**Attractivity of Equilibrium Graph Colorings**

Now, even if we know all possible invariant graph colorings, one could wonder: how can we know what the evolution of an arbitrary graph coloring is, when it is not invariant? The answer is given by showing that any initial (non-invariant) graph coloring will necessarily converge or be attracted by some equilibrium (invariant) graph coloring, as we are going to see in the sequel.

[Should we use Markov chains – like finite automata – at least for $k$-regular graphs???
In other words, given any $k$-regular graph (with $k$ odd), can we find ALL the invariant colorings – other than the trivial ones – and then say whether an arbitrary initial coloring will converge to an invariant one and which one???]

## References


Arthur, W.B. (1989). Competing technologies, increasing returns, and lock-in by historical small events. *Economic Journal*, vol. 99, pp. 116-131.

Bailey, N.T.J. (1975). *The Mathematical Theory of Infectious Diseases and Its Applications*. New York: Hafner Press.

Brémaud, P. (1999). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. New York: Springer-Verlag.

Davis, J.A. (1967). Clustering and structural balance in graphs. *Human Relations*, vol. 20, pp. 181-187.

Durrett, R. (1981). An introduction to infinite particle systems. *Stochastic Processes and Their Applications*, vol. 11, pp. 109-150.

Griffeath, D. (1979). *Additive and Cancellative Interacting Particle Systems*. Berlin: Springer-Verlag.

Harris, T. (1974). Contact interactions on a lattice. *Annals of Probability*, vol. 2, pp. 969-988.

Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology*, vol. 21, pp. 107-112.

Holley, R.A., & Liggett, T.M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Annals of Probability*, vol. 3, pp. 643-663.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, vol. 83, no. 6, pp. 1420-1443.

Keeling, M.J. (1999). The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London*, *Series B*, vol. 266, pp. 175-190.

Kindermann, R.P., & Snell, J.L. (1980). On the relation between Markov random fields and social networks. *Journal of Mathematical Sociology*, vol. 7, pp. 1-13.

Liggett, T.M. (1985). *Interacting Particle Systems*. Berlin: Springer-Verlag.

Morris, S. (2000). Contagion. *Review of Economic Studies*, vol. 67, pp. 57-78.

Prun, B., & Fort, J.C. (1991). *Stochastic Processes on a Lattice and Gibbs Measures*. Dordrecht, Boston & London: Kluwer Academic Publishers.

Spitzer, F. (1970). Interaction of Markov processes. *Advances in Mathematics*, vol. 5, pp. 246-290.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis. Methods and Applications*. Cambridge, New York & Melbourne: Cambridge University Press.

Watts, D.J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 5766-5771.

## Appendix

1.

The **adjacency matrix** $A$, corresponding to the graph $G$, is the $N \times N$ symmetric matrix $A = \{a_{ij}\}$ (without diagonal elements) such that $a_{ij} = 1$, if $j \in V(i)$, $a_{ij} = 0$, otherwise.

Just in the way the degree of a single vertex was defined, we may define the **degree of a set of vertices** $U \subseteq V$, denoted as deg($U$), as the total number of all links connecting the vertices of $U$ with any other vertices of $V$, i.e., $\deg(U) = \sum_{i \in U} \deg(i) = \sum_{i \in U} |V(i)| = |V(U)|$, where $V(U) = \bigcup_{i \in U} V(i)$ is the set of all neighbors of $U$. Since part of the neighbors of the vertices $i \in U$ may lie inside $U$ and part of them outside $U$, deg($U$) may decompose as follows:

$$k(U) = \deg_{in}(U) + \deg_{out}(U),$$

where $\deg_{in}(U)$ denotes the number of links connecting vertices in $U$ with other vertices of $U$ and $\deg_{out}(U)$ is the number of links connecting vertices in $U$ with vertices outside $U$.

3.

A ***m*-coloring** of the graph $G$ (which is then called 'colored graph') is a labeling $f$: $V \rightarrow S$, where the set of labels $S$ is a finite set ($|S| = m \geq 2$) composed of $m$ colors 0, 1, …, $m$ - 1. When $f(i) = s \in S$, we say that "the color of $i$ is $s$" or that "$i$ is colored $s$." A set of vertices $U \subseteq V$ is called **monochromatic** if $f(i) = s$, for all $i \in U$, and **polychromatic** if there exist $i, j \in U$, $i \neq j$, with $f(i) \neq f(j)$. In case of **binary** or **2-colorings**, i.e., when $S = \{0,1\}$, a polychromatic set becomes **dichromatic**. Similarly, we can define monochromatic and polychromatic or dichromatic links.

4.

Furthermore, for any set of vertices $U \subseteq V$, we set $V^s(U) = \bigcup_{i\in U} V^s(i)$ for the set of all $s$-colored neighbors of $U$.

5.

**Lemma 1.** For any color $s \in S$,

$$\sum_{i\in V} n^s(i) = \deg(V^s) .$$

**Proof:** Fix $i$ and define the characteristic function $\chi^s(i)$ as $\chi^s(i) = 1$, if $f(i) = s$, $\chi^s(i) = 0$, otherwise. Then, using the adjacency matrix $A = \{a_{ij}\}$,

$$\sum_{i\in V} n^s(i) = \sum_{i\in V}\sum_{j\in V} a_{ij}\chi^s(j) = \sum_{j\in V}\left(\sum_{i\in V} a_{ij}\right)\chi^s(j) = \sum_{j\in V}\deg(j)\chi^s(j) = \sum_{i\in V^s}\deg(i) = \deg(V^s),$$

since A is symmetric and $\sum_{j\in V} a_{ij} = \deg(i)$. ■

The previous relation is simplified in the case that $G$ is a ***k*-regular** graph, which means that all vertices have the same degree, i.e., $\deg(i) = k$, a positive integer, for all $i \in V$.

**Corollary 1.** If $G$ is $k$-regular, then:

$$\sum_{i\in V} n^s(i) = k\, N^s, \text{ for all } s \in S,$$

$$\sum_{s\in S}\sum_{i\in V} n^s(i) = k\, N.$$

Similarly, counting all the $s$-colored neighbors of an arbitrary set of vertices, we obtain:

**Lemma 2.** For any set of vertices $U \subseteq V$ and any color $s \in S$,

$$\sum_{i \in U} n^s(i) = \deg_{in}(V^s(U)) + \deg_{out}(V^s(U)) \text{.}$$