The Last Bottlenecks in Communications

Eli Noam

Columbia University

First Ithiel de Sola Pool Memorial Lecture

Delivered at M.I.T., Department of Political Science

October, 27 1990

I. Introduction

Ithiel Pool provided intellectual blasts of major energy throughout his academic career. In the policy field, his intellectual contribution was especially important to the question whether constitutional rights of free speech apply also to electronic media. Pool championed freedom from governmental restrictions in the provisions of content and conduits. On the whole, his ideas have made headway. In America and increasingly abroad, electronic highways proliferate, as do the providers of electronic and video information. Electronic networks thus appear well on their way towards openness and freedom.

But are they really? Just because one set of restriction disappear does not mean that new and perhaps unanticipated bottlenecks do not emerge instead. And this is the subject of my discussion. I will identify two major new bottlenecks. The first is, paradoxically, the result of the exercise of one of the freedoms which Pool advocated – the freedom of association. I will argue that the cumulative impact of this freedom in the telecommunications field may well lead to restrictions in the exercise of speech.

The second bottleneck is very different from the first. Even if all media channels become legally free and technically powerful in providing information, they still have to deal with the increasing inability of individuals and institutions to deal with these information flows. This bottleneck, which I call the problem

of the "last twenty inches", raises challenges to organizations and technologists to conceive systems and devices that can manage and screen information, or else information flows will be backed up rather than open.

To understand why freedom of association may lead to reduced freedom of specch, we have to understand more about how group communications form and why they disintegrate. This will be the subject of the more theoretical discussion of the next three sections.

II. THE CONCEPT OF NETWORK

Let us first understand what networks are, because they are a key concept in the communications field. But not just there. Networks abound as physical facilities or as relational systems (such as those of "old boys" or political supporters).

The term "network" goes a long way back; it is used, in Exodus by the Supreme Regulator herself: "And God spoke unto Moses, saying..... You shall also make it a grating, a network of brass..." XXVII, V. 4. In Hebrew, the word is "reshet," (net) similarly used today for telecommunications and other networks.

The term is used by most academic disciplines, and with a variety of meanings. Chemists apply it to arrangements of molecules.[1] Biologists to cell structures.[2] Mathematicians to

---

[1] Zacharisen, W. M., "The Atomic Arrangement in Glass," October 5, 1932, Journal of the American Chemical Society, Vol. 54, No. 10, p. 38-42, Washington, D.C.

topology.[3] Electrical engineers to distribution systems (for high voltage), or for circuit configurations of components (for weak voltage).[4]

Operations researchers use a network terminology to solve shortest path problems, maximum flow models, and optimal routing.[5] Computer scientists apply the term for computer interconnections in hardware, and to implementation algorithms in software.[6]

In the social sciences, political scientists use the concept of networks in discussing, for example, hierarchies, interactions, gatekeepers, and policy communities.[7] Sociologists and social

---

2 Knox, Robert, 1830, Elements of General Anatomy (translated from Inst. edition of "Beclard's Anatomy" by D. A. Beclard, (Edinburgh, Scotland: Maclachlan & Stewart), p. 214.

3 Klingman, David J. and Mulvey, J. eds., Network Models and Associated Applications, Amsterdam and New York: Elsevier North, Holland, 1981.

4 Karni, Shlomo, 1986, An Analysis of Electrical Networks, New York: Wiley, pp. 1-4.

5 Elmaghraby, Salah E., 1970, Some Network Models in Management Science, New York: Springer, pp. 1-3.

6 Klingman, David J. and Mulvey, J. eds., Network Models and Associated Applications, Amsterdam and New York: Elsevier North, Holland, 1981.

7 Richardson, Jeremy John, Gunnel, Gustafson and Art Jordan, as cited in Rhodes, R.A.W., Power Dependence, Policy Communities and Intergovernmental Networks, Colchester, Essex; Department of Government, University of Essex, Wivenhoe Park, 1985, pp. 6-8.

anthropologists[8] speak of network dyads -- interpersonal linkage between two persons in which each is indebted to the other, and similar in some ways to the exchange relation of economics.

Among the social science disciplines, economists have probably paid the least attention to networks. Closest are public choice theories of group formation.[9]

One should not look at telecommunications as a technical facility. The network structure, if left to its own devices, is a reflection of the underlying group interaction in society.

In that sense, a telecommunications network is merely one example for the conflicting forces of integration and centrifugalism common to many social processes. Wherever one looks, people break up all kinds of social networks of interaction and form new ones.

Examples abound. They include, for the United States, public education, mass transit, public safety, dispute resolution, pension systems, health services, electrical power distribution, stock

---

8 Barnes, J.A., 1954, "Class and Committees in a Norwegian Island Parish," Human Relations, Vol. 7, Lazarsfeld, Paul and Merton, Robert, 1954.
Bott, E., 1957, Family and Social Network, London, 2nd Ed., 1971.
Boissevain, J., 1979, "Network Analysis: A Reappraisal," Current Anthropology, Vol. 20.

9 Noam, Eli, "The Next Stage in Telecommunications Evolution: The Pluralistic Network," Paper presented at the ITS Conference, MIT, June 1988, Working Paper #316, Columbia University CTIS.
Heal, Geoffrey, "The Economics of Networks," Columbia University, unpublished paper, 1989.
Economedis, Nicholas, "Desirability of Compatibility in the Absence of Network Externalities," American Economic Review, Vol. 79, No. 5, pp. 1165-1181, (Dec. 1989).

exchanges, department stores, or universities. And telecommunications.

III. THEORIES FOR THE FALL OF NETWORKS

In telecommunications, forces of change are transforming the traditional system of monopoly. A number of explanations have been offered for its demise. There are three major types of theories.

1. Technological Explanation

"More powerful technology, and the merging of telecommunications and computing, lead to diversity and break down traditional barriers."

This view is influential in an engineering-oriented industry such as telecommunications. But it does not provide a sufficient condition. After all, the same transmission and switching technologies have been available almost anywhere on the globe, but very few countries have left the traditional system.

2. Political Explanations

> "In the information age, a network monopoly becomes too powerful and uncontrollable and its scope needs to be limited."

The problem with this view is that the introduction of diversity is only one policy option out of several. An alternative response to political power or regulatory inefficiency might well

be a stricter regulation, or nationalization, not necessarily of freedom to enter.

3. Market Structure Explanations

"Monopoly's inefficiency leads to the emergence of diversity."

This view has holds a premise of inefficiency of the traditional system. A multi-carrier market structure is believed to be emerging due to some inadequacy of the traditional system. Yet, if inefficiency were the causal force for rival entry, Egypt or Mexico, for example, should have introduced competition long before the U.S. and Japan, which had arguably the most advanced and ubiquitous networks in the world even before embarking on their liberalizing policies. This approach, too, does not provide a full answer.

4. A Fourth Approach: The Tipping of Network Coalitions

In contrast, I would like to advance an alternative view, that of the dynamics of group formation. It is the approach of public choice and club theory.

> "Political group dynamics lead to an overexpansion, and to incentives to exit from a sharing coalition, and to an eventual 'tipping' of the network from a stable single coalition to a system of separate sub-coalitions."

This view of success undermining its own foundations is, from the monopoly's perspective, deeply pessimistic, because it implies that the harder their efforts and the greater their success, the closer the end to their special status is at hand. Like in a Greek tragedy, their preventive actions only assure their doom.

III. A MODEL OF NETWORK TIPPING

Perhaps the best way to look at a network is as a cost sharing arrangement between several users. Fixed costs are high, marginal costs low, and a new participant C helps the incumbents A and B to lower their cost. This can be illustrated with a simple model. (Graph 1) The horizontal axis is the size of the network. The vertical axis are costs and benefits. We assume, for the moment, that the users share costs equally. Average costs will initially come down but, later rise, hence the U-shape of the cost curve. [10] Secondly, an individual's utility is given by the benefit curve, which is rising because we assume network externalities to exist, that is, that a network member gets benefits from other members' presence.

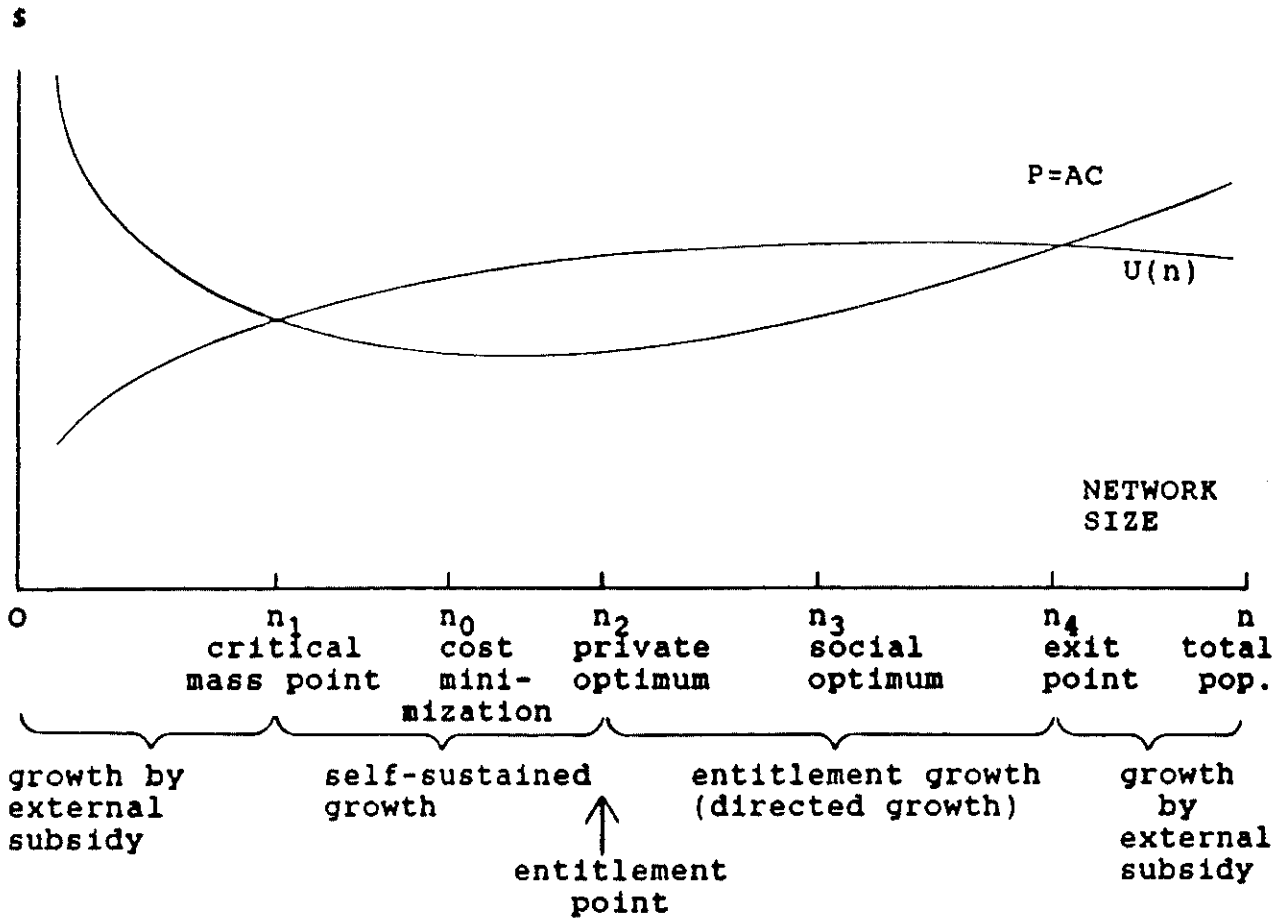
Graph 1 about here

---

[10] As the network reaches universality, connecting the last members increases cost. An indicator for rising costs: In the Bell System, the average capital investment cost per new telephone grew steadily (in 1982/3 dollars.)

1945:	$1928
1955:	$2050
1965:	$2580
1975:	$3960
1985:	$4624

FIGURE 1



STAGES IN NETWORK EXPANSION

What we then see is, first, that below a "critical mass" point, a network will not be feasible, unless supported by a subsidy of sorts. The club is too small.

Second, in early phase of network growth, the early users can lower their cost by adding members. Beyond the critical mass point, self-sustained growth takes place.

Left to itself, the network association will grow to n_2. Applicants beyond that point raise average cost, and add fewer externality benefits. The club will not accept new members. But from a societal point of view the optimal network size in an equal price system diverge from the private optimum. Social welfare still increases at n_2 because the benefits to new entrants is factored in, not just to the incumbent.

This then usually leads to politically directed growth beyond private optimum n_2, which can be termed an "entitlement growth" because it is based on political arguments of <u>rights of equal treatment</u> to participate in the network. We normally call this the universal service obligation.

What is important is that universal service need not be something imposed externally by a government. One can show that the <u>internal</u> dynamics of network members will also take the network toward universal service, and toward its own disintegration. As an illustration, see the following cost figures for incremental wiring.

I mentioned that a network will cease to grow on its own after private optimum n_2. But this conclusion was based on a pricing

scheme of equal cost shares. Yet there is no reason why such equality of cost shares would exist in the real world. Freedom of association leads to the formation of groupings among the network users. If there is a decision mechanism that permits the majority of network users to exercise control over pricing, they will impose higher cost shares on the minority.

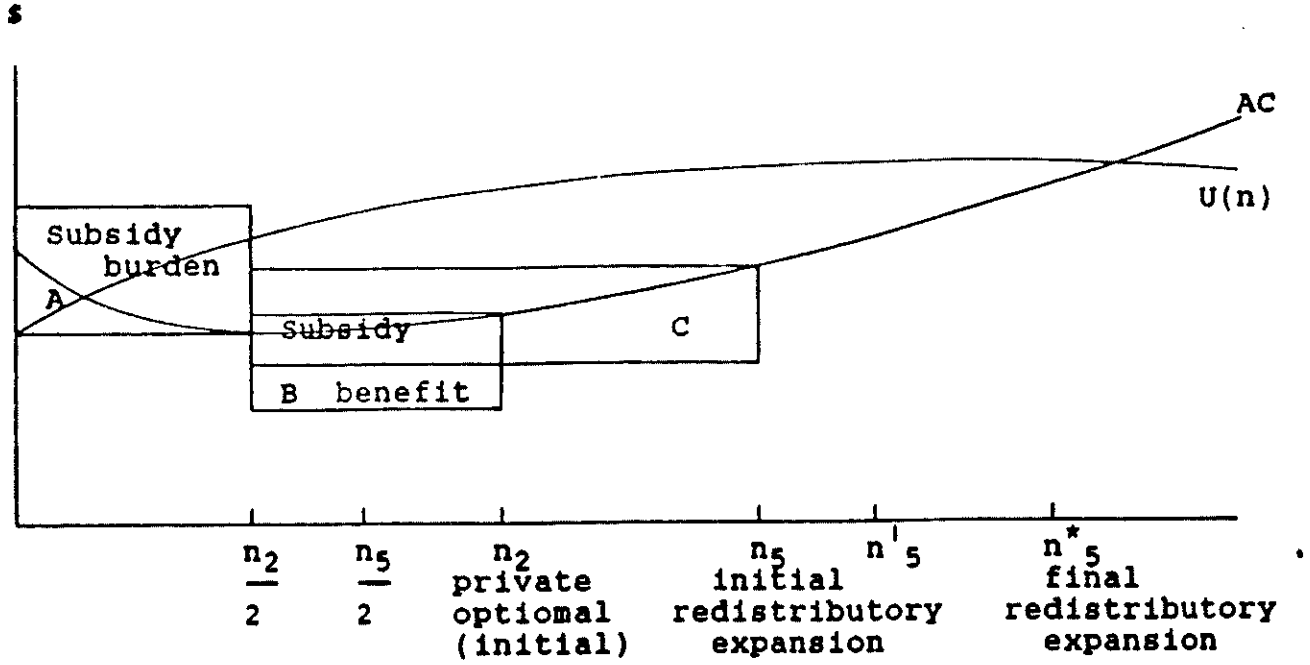
This can be analyzed along the lines of public choice modes with minimal winning coalitions. A majority will establish itself such that it will benefit maximally from the minority. The minority that can be maximally burdened are the users with less elastic demand for telephone service. But there is a limit to the burden. If price gets pushed too high, subscribers will drop off. Hence, the majority will burden the minority with a price up to the latter's positive utility.

Graph 2 shows this relation. The benefit function gives a subscribers's utility for various network sizes n. Area A constitutes the subsidy to the minority; B is the benefit of the subsidy to the majority.

GRAPH 2 ABOUT HERE

But such redistribution is not a stable equilibrium. Before, network size was n_2. But with internal redistribution, several things happen. The majority enjoys the added utility from added network members, while most of its cost is borne by the minority. The majority will therefore seek expansion, which will get it up to a size of n_5. The subsidy now gets spread over more people, up to a point, where marginal externality benefits are equal to

FIGURE 2



REDISTRIBUTION IN NETWORKS

marginal cost, including the dilution of subsidy.

And this is not the end of the story. With expansion to n_5, the majority is now larger than before, and it can also tax a larger minority than before. Hence, the expansion process would take place again. This process would continue, until an equilibrium such as n_6 is reached. (We can model that point relatively simply).

As this process of expansion takes place, the minority is growing, too, and its size can increase beyond critical mass. Minority coalition members find themselves better off associating in a new network, and abandon the old one if they can do so legally. Barriers to exit become lower with the increasing size of the users who exit, and with the lower cost of technology. The rule of thumb is that the largest 3% of users account for 50% of network. As the absolute size of the communication of these users increases and the cost of the network technology comes down, it takes less and less parties to start a new club.[11]

This process of unravelling of the network starts even earlier if a new network has the right to <u>interconnect</u> into the old one.

---

[11] Strictly speaking, one could relax the requirement of a U-shaped cost curve. The analysis could also be used for continually falling average costs. Here, an exit point may be reached if network externalities, after a point, become negative and downward sloping, e.g., because the crowding of the network reduces its usefulness. The analysis can also be applied even without negative externalities, and in the presence of continuing by falling average costs. Then, a network will reach the size N of the population. But with the political coalition formation taking place, the redistribution will still burden N/2 with a subsidy, leading to its exit if $N/2>n_1$, the critical mass point.

Then its members can have their cake and eat it, too. They can enjoy the externality benefits of a larger reach while not being subject to the redistributory burden.

But there is a social cost to this. A private start-up investment in a network before take-off is based on an expectation of eventual break-even and subsequent positive net benefits to members. An additional new network could latch on that critical mass. Hence, a loss would be entirely borne by the initial network participants while the benefits would be shared with other entrants. The implication is that it pays to be second. A situation of market failure exists. Innovation is slowed.

And in such a situation, there can be a role for direct outside support, such as by a government subsidy as a catalyst role for the various and disparate parts of the network.

To conclude: The model thus shows that a network coalition, left to itself under majority-rule principles, expands beyond the size that would hold together. Such an arrangement can be stable only as long as arbitrage is prevented, or as long as the minority cannot exercise political power in other ways, and, most importantly, as long as it has no other choice.

Otherwise, under the assumptions of the model, the network progresses beyond its "tipping point." Its cohesion breaks up and a pluralistic multi-network system emerges. The success of communalism in expansion of telecommunications creates the forces

for particularism. [12]

IV. THE NEW GROUP NETWORKS

The theoretical discussion of the previous section aimed at demonstrating the dynamic of disaggregation in networks. If one gives individuals the freedom of association, they will form new types of interlinkages which we call networks. What are some of the long-term implications?

1. Networks will become transnational

As the cost of transmission continues to drop, the network associations will not be territorially organized. Territoriality was based on the need for a network architecture that primarily minimized cost by minimized transmission distance. It led to the creation of the 'German network,' or the 'French network.' This technological and economic territoriality suited governments everywhere just fine, because they, too, were based on territoriality of jurisdiction, and could thus conveniently exercise control and even ownership over "their" network. But things are changing. Now, networks are increasingly becoming pluralistic group affairs. Groups break off parts of their

---

[12] And this process of network evolution in the network is expansionary internationally in the sense that one country's developments affect the others. The more interrelated countries and economic activities are, the less likely are there stable solutions to separate policies. And where instabilities exist in one country, they ripple through the entire international system.

In the past, international regimes have often been used to stabilize domestic arrangements. But the symmetrical scenario is being played out now in the opposite direction, as international trends undermine domestic stability.

communications needs from the public network and aggregate them in their own associations. Banks, insurance agencies, airlines, automobile manufacturers, and many others communicate with each other on increasingly specialized networks. Advertising agencies, marketers, printers, and media do so similarly. Another example is automobile manufacturers, their suppliers, dealers, and financiers.

Territoriality becomes secondary. Many of these communities of interest transcend national frontiers. Their interests are continental and global, and so are their networks. When the computers of brokers and investment banks in New York are interconnected by a continuous network and interact with those in Tokyo and London to trade and clear transactions, one cannot say anymore that there is a New York or Tokyo market. There is no physical locus for the market anymore. **The network becomes the market.** Transactions are not conducted at any particular physical point.

2. New Electronic Neighborhoods will Emerge

A few years ago, it became fashionable to speak of communications creating the "global village". There was something inspiring in this image, communal and peaceful. But there is nothing village-like in the unfolding reality. Instead, groups with shared economic interests are extending national group pluralism through the opportunity to create global interconnection with each other into the international sphere. Indeed, communications make international pluralism easier because it is easier to reach

critical mass for subnetworks if one aggregates across several countries.

The new group network do not create a global village, they create instead the world as a series of electronic neighborhoods. In the past, neighborhoods had economic and social functions. In New York for example, there are Chinatown, the Garment District, Wall Street, Madison Ave., or the Theater District. Elsewhere, there are regions with specialized production. Solingen and Sheffield for cutlery; Lyons for silk; Hollywood for films; Silicon Valley and Route 128 for microelectronics.[13] Production clusters create economies of aggregation that substitute for the economies of scale and scope of the giant multi-product firm. Physical proximity was a key. But now, group networks can serve many of the functions of physical proximity. They interconnect specialized producers, suppliers, buyers, experts, and markets. They create new ways of clustering, spread around the world.

Some of these electronic neighborhoods will be nicer than others. They will perform better, faster, and often even cheaper. In developing countries, the networks of those transacting with the world are already becoming better than those of local people. In places like China of Egypt, a two-tier communications systems has in effect emerged.

Networks might also be stratified along socio-demographic dimensions. Already, some long-distance resellers in the U.S. offer

---

[13] Piore, Michael & Charles Sabel The Second Industrial Divide: Possibilities for Prosperity New York: Basic Books, 1984

bonuses to churches if they sign up their members. Such marketing efforts can lead over time to identification of some network with particular ethnic, religious, or political groups. Similarly, some networks may be shunned by labor union members if they have a history of labor problems.

People or businesses could become identified with "their" network. A few months ago, New york Telephone proposed splitting the 212 area code, with the Bronx and parts of Manhattan get a new code. Many Manhattanites were upset to be lumped together with the Bronx and made themselves heard.[14] Governments might try to maintain system of internal redistribution by resorting to taxation and allocation. A value-added tax on communications would be a sensible substitute for the present hidden system. But it will not be easy to define what will be taxed, or to measure it, or to prevent the taxed electronic flows simply to bypass the jurisdiction.

3. Networks will assume political power as quasi-jurisdictions

Historically, the nation state was at tension with cross-border allegiances--whether proletarian international solidarity, rebellious youth culture, international financial capital, or ethnic minorities. The new network environment weakens national cohesion. It strengthens particularism and internationalizes it. It is difficult for a state to extend its powers beyond traditional frontiers, but it is easy for the new networks to do so.

---

[14] The alternative proposal was to assign all new lines to the new code, which would have created the possibility, for the first time anywhere, for "his-and-her" area codes.

Furthermore, these network associations possess and acquire powers of their own. They already may link powerful entities, and can bring their combined powers to bear. For example, the combined weight of the members of the SWIFT banking network got the powerful national PTT monopolies to cave in on a number of crucial issues. And there is no reason to expect the power of network combinations to be directed only at communications issues. Once groups are in constant touch, they may as well get organized on other issues, too. **The communications network becomes the political network.**

They will coordinate in the economic sphere. When it comes to the role of information, the line between competition and cartel coordination has always been a fine one. In the 1920s, various American industries established so-called fair-price bureaus that gave each member of the industry a convenient look at what its competitors were charging. This practice was outlawed in a series of antitrust cases. Imagine if one leaves instead information exchange to a series of artificial intelligence programs communicating internationally. One has a real problem of conceptualizing, detecting, and preventing international cartels. One person's collusion is another person's programmed trading. **The network becomes the cartel.**

The network associations are also likely to become quasi-jurisdictions themselves. They have to mediate the conflicting interests of their members. They have to establish cost shares, sometimes creating their own de-facto taxing mechanism as well as redistribution. They have to determine major investments, to set

standards, to decide whom to admit, and whom to expel. As a network becomes more important and complex, control over its management becomes fought over. Elections may take place. Constitutions, bylaws and regulations are passed. Arbitration mechanisms are set up. Financial assessment of members takes place. **Networks become political entities.**

Thus, we may be witnessing the creation of new and often extraterritorial forms of new quasi-jurisdictions that are not clearly subordinated to others. In response, governments might create forms of domestic and international regulatory mechanisms for specified sets of problems, possibly based on global networks themselves that continuously collect and exchange information, track activities, and coordinate enforcement.[15]

4. Networks will exercise power toward their members

Perhaps, the major question is whether a network group can dominate its own members, or be restrictive in its permission of other to join. The power of the network becomes most obvious when it is operated by a dominant entity.

Examples:

*The network of a university such as Columbia or MIT can be quite restrictive toward its members. It can and does limit terminal equipment and options, charge monopolistic prices, and it

---

[15] The optimal size of jurisdictions was always dependent on communications. French departments were based on the distance that a horseback rider could cover in a day. Transportation and communications technology changes the optimal size. It's hard to imagine a voluntary European integration without telecommunications.

could legally refuse to serve political activist groups.

*The major U.S. videotex service, Prodigy, prevents its users groups discussing politics on the system as well as the Prodigy system itself. When Prodigy, which provides extensive messaging service, announced that it would raise the rates for such messages, a group of subscribers posted notices in a "public area" of the system encouraging other subscribers to protest. When Prodigy removed these messages, the protesters turned to the private message feature, and sought help from advertisers. Thereupon, Prodigy cancelled the subscriptions of the protesters.[16]

*The National Science Foundation recently urged NSF sites to remove from computers networks scanned image files of arguably pornographic images. (M. Rotenberg, communication)

*In 1987 a debate raged at Stanford University over a joke file on the University's computer system. Because it contained jokes offensive to some groups, the university was pressed to impose restrictions on content.

*On the public networks, too, content control emerges. Telephone companies recently sought to establish their right to restrict otherwise lawful communications if it was harmful to their image. [17]

*Employers frequently block the ability of their employees to reach certain numbers. While this is based on protections against

---

[16]Professor Henry Niman, per Marc Rotenberg, communication

[17] Some regulatory commissioners, most notably New York (but not the FCC) resisted; but at least one major court decision, in a muddled opinion, ( ) seemed to permit restriction.

running up telephone bills generated by dial-it services, the principle could be extended to an exclusion of messages of a type undesirable to employers, such as those of labor unions.

*In so-called intelligent buildings, landlords provide communications to occupants. These "shared-tenant services" are largely under the control of the building owners, whose interconnection decisions determine which networks tenants can reach.

Petty monopolies can thus emerge, largely unencumbered by the protections built into the public network, at least in the past, by law, custom, and regulation. The option is exit, which in a university setting may mean giving up tenure and departuring to another institution.

Are there freedom of speech rights for users (in network terminology "common carriage obligations") in group networks? The scope of these rights is undefined. Constitutional First Amendment rights do not appear to exist, given the absence of state action. Regulatory impositions of such obligations are possible, but are limited by the rights of groups to substantially define their membership and the rules under which they operate, especially where a major purpose of the groups is communication, and thus the exercise of a fundamental right itself, i.e., of speech. In such circumstances group activities have protection from restrictive regulation. In other contexts, the exercise of speech rights is stymied by access problems, especially to the workplace or to the shopping malls that take today the role of public gathering spaces.

By analogy, the access to networks might be foreclosed, and with it its free speech potential.

Even where network groups are organized democratically, they may well be restrictive. A major function of liberties, after all, is to protect minorities from unsympathetic majorities. In the public sphere, guarantees of free speech against governments are part of constitutions. In the network environment, the granting of access and non-discriminatory content-neutrality is required of the general "public" networks by law or common carriage regulation. But common carriage does not necessarily apply to group networks. Groups may institute restrictions on the exercise of speech over their network, and assert that their status is alike to publishers, with no rights of users. They can exclude certain subjects from being discussed, or certain speakers from having access to the network. This could become particularly an issue when telecommunications networks gain the ability to transmit video programs. It is true that individuals could form alternative networks if they are being restricted. Thus, market forces could help, but not if some of the networks control some segments of a chain of communications, or where the ability of any link in such a chain to institute content-based tests would impose transaction costs on the entire system. It is for similar reasons that society has adopted the use of the legal tender and of commercial paper to permit low-cost transactions. Common carriage has a similar rationale.

It will not be desirable or possible to extend the common

carriage model all the way into the last small group network or into a broadcast-like one-way network. At the same time, it is not supportive of free speech to let sub-networks to set restrictive policies on content, just because they are less inclusive than the public network. A three-tier system may therefore be helpful:

1. **Public networks:** operating as common carriers.

2. **Group networks holding themselves open to the public or possessing bottleneck powers:** also common carriers.

3. **Closed and small user groups:** can set their own rules.

Group formation always had a double-edged aspect. On the one hand, they were an extension of individual rights. De Tocqueville noted that the "right of association . . . almost is unalienable in its nature as the rights of personal liberty." On the other hand, freedom of association led to situations inimical to individual as well as to a more general public interest. While many are agreed with the significance of pluralism,[18] others note the negatives.[19]

---

[18] Dahl, Robert A. Who Governs?: Democracy and Power in an American City, New Haven: Yale University Press, 1961

Lasky, C. Foundations of Sovereignty, 1921

Lindblom, Charles E. The Intelligence of Democracy: decision making through mutual adjustment, New York, Free Press, 1965

Truman, David B. The Governmental Process: Political interests and public opinions New York City, Knopf, 1951

[19] Nisbet, Robert A. The Quest for Community: A Study in the ethics of order and freedom, New York, Oxford University Press, 1953

Lowi, Theodore J. The End of Liberalism: the second republic of the U.S., 2nd Ed. New York, Norton, 1979

In this context, the exercise of freedom of association may lead to group formation that are restrictive of speech. Hence, the evolving pluralistic structure of telecommunications may bear the seeds for a new type of bottleneck to the free flow of information that did not exist on the traditional public network and its common carriage.

V. Personal Networks and Information Overload

But why stop at networks for groups? If the trend is from national public networks covering the entire population to a pluralist system, why not expect still further disaggregation? This additional step means individualized networks, or personal networks, which may be called PNs, analogous to PCs. Before dismissing the notion of PNs as extravagant, remember that twenty years ago nobody expected personal computers PCs, and nobody expected computers to end up on everybody's lap, either.

What does a personal network mean? It means an individually tailored network arrangement that fits an individual's communications needs. It does not necessarily mean a separate physical system, except for inside wiring and maybe the last mile of circuits, plus some radio-mobile links, and terminal equipment. The rest consists of what is called virtual networks, provided by a whole range of service providers and carriers, not just one, and packaged together to provide easy access to an individual's primary

---

Kariel, Henry S. The Decline of American Pluralism, Stanford, Calif. Stanford Univ. Press, 1961

communications needs: friends and family; work colleagues; frequent business contacts, both domestic and foreign; data sources; transaction programs; and video publishers frequently accessed; telemetry services such as alarm companies; bulletin boards scanned etc. Contact to and from these destinations would move with the individuals, whether they are at home, at the office, or moving about.

VI. The Human Bottleneck

As these changes are taking place, what used to be called the public network becomes an incredibly complex overlay of multiple subnetworks of various kinds. It becomes a fiber "ether." **Information flows become the environment.** More information, faster, cheaper. But where is this taking us? When all is said and done, the real challenge for the future technology may not be how to transport information, but rather how to deal with it once it is there. And this gets me to the second long-term bottleneck of communications.

Sometimes the worst is to get what one wants. And perhaps this is happening to us with the revolution in information and communications. The technical trends are toward abolishing all bottlenecks. Except one.

This last bottleneck are the last 20 inches, the 20 inch distance from the display terminal to the human brain. The human sensors and processors, eye, ear, and brain can only handle so much information. They are subject to biological constraints.

A tidal wave of information is flooding society. There are

more books written than ever, and probably less books read than ever. Cable TV provides dozens of new channels of TV. Soon, no doubt, voice recognition technology will finally reach the state that any random thought could be typed as one speaks, and almost instantaneously distributed by electronic mail and broadcast-fax to hundreds of innocent bystanders on our personal network.

Can we quantify the information trend? In the first half century after Gutenberg, about 20 mil books were produced. That is not a small number. But it is dwarfed by today's figure by a factor of about 50,000. In the U.S. alone, about 2.3 billion books were distributed in 1987. Per capita paper consumption in the US has increased in the past decade by 66% to about 50 pages per person per day. It has been said that 80 or 90% of all scientists who ever worked are alive today. That's the good news. The bad news is that it's a similar figure for lawyers, too. In this area too, Ithiel Pool did pioneering research in quantifying the trends, together with his American and Japanese collaborators such as Inose, Takasaki, Hurwitz, and Neuman.[20] They quantified media by using the common denominator of actual words which they contained, and found that in 1960 the mass media supplied to an average American household was about 3 million words per day, including unwatched TV, unread papers, unlistened to radio, etc. By 1980,

---

[20]Pool, Ithiel de Sola, Inose, H. , Takasaki, N. and Hurwitz, R., 1984 Communications Flows: A Census in the United States and Japan, Tokyo: University of Tokyo and Elsevier Press.

Pool, Ithiel de Sola, and Neuman, W.Russel, 1986, "The Flow of Communications into the Home in S. Ball-Rokeach and M. Cantor, eds. Media, Audience, and Social Structure. Sage

this figure had increased by 267% to 11 million words.

Graph 3 about here

Obviously, only a tiny fraction of these media words that reached the average household is actually consumed, about 60,000 media words/day, which still comes to a pretty extraordinary 1 word per second. This number was up by 51% from 1960 to 1980. [21]

When one starts factoring in the price per word of different media, one finds that broadcast words are a real bargain for consumers relative to print words, which explains why their share increases while attempts are made to charge more for them than before, such as subscription cable and pay-TV.

Given the steady increase in information supplied to us, the real issue for future technology does not appear to be production of information, and certainly not transmission, but rather absorption.

Several strategies are possible to increase absorption:

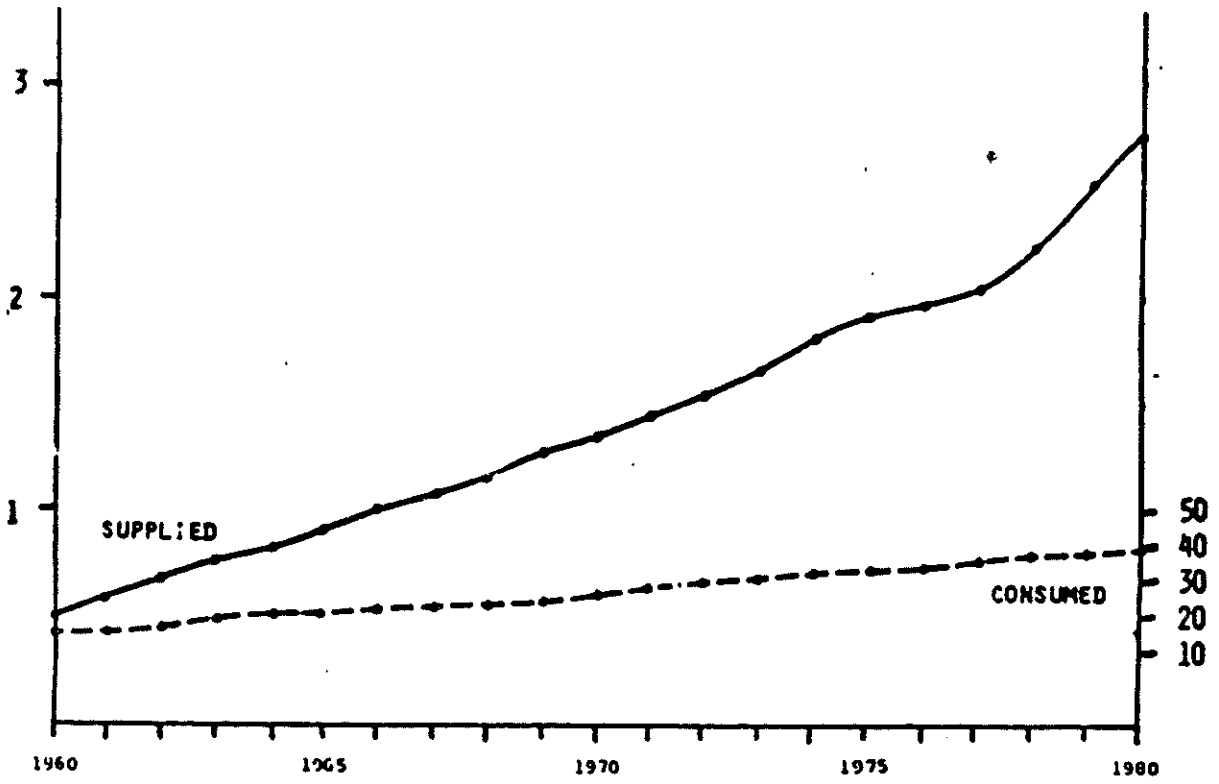
1. **Education:** i.e., make humans smarter. But there are severe limits to this, as one finds out after about two semesters of teaching experience.

2. **Add time allocation:** i.e., spend more <u>time</u> on informational activities. That is clearly happening. The average cable TV household has its set on for an extraordinary 8 1/3 hours per day, which is 2 hours more than a household without cable.

---

[21] TV, incidentally, accounted for 2/3 of word consumption, and this does not even count the visual images that are not part of the analysis. If a picture is counted as a thousand words, the increase of information would be truly dramatic.

TELEVISION/CABLE



Individuals create coping strategies, such as scanning correspondence while answering a telephone call while listening to radio news. In office settings, people spend more time on paper flow; lunches get shorter, work hours longer. Obviously, there are limits to this strategy.

3. **Tinkering with Mother Nature**, by pharmacological or biological engineering. This is not an attractive proposition. Probably and hopefully it has natural and ethical limits. One could even bypass eyes and ears and get directly into the brain, though one shudders at the totalitarian potential.

4. **Substitute information storage for human reception.** We all have ever-increasing piles and files of things to read. This produces a temporary illusion of a match between inflow and outflow. It leads to technologies and training that emphasize search skills over knowledge.

5. **Change the way information gets presented.** Eyes can get visual information at a broadband megabyte-rate. If the TV action is too slow, one gets bored. Written information gets absorbed at a much slower rate, about 200 bits per second. Ears are even slower, about 150 bits per second. Which is one reason why in an information-intensive environment, professional person-to-person conversation declines. And the tactile sense can get perhaps 20 words/min., or about 15 bps, using Braille.

Thus, visual information is by far and away the fastest, if it uses the entire bandwidth of the eye's ability. But, importantly, print language cannot do that. Print takes up only

a tiny fraction of our absorptive capacity. We are using hopelessly outmoded Phoenician and Latin communications protocols. But we are stuck with them. The written word is often sacrosanct. Try to change a letter in the Bible, and you start a religious war. The form of written language has hardly changed in centuries, and society has a big investment in this particular form of standardization. Society needs compatibility, and symbols are part of the infrastructure. Social and cultural fabric revolves around them. Therefore, even streamlining the needlessly complicated spelling of the English language would be a culturally traumatic event, and unlikely to happen outside a tiny circle of professionally eccentric poets.

So instead of junking the Latin alphabet and traditional forms of written language, what is more likely to happen is a shift to a multimedia form of communications with more visual and symbolic information.

Television advertisements are an example. It is easy to make fun of them, but they pack a lot into 30 seconds of picture, voice, music, and written language, all superimposed on each other, and conveying messages on a variety of levels. Another example are speech presentations, with their increasingly elaborate visual aids, now even including video clips. There are millions of camcorders in search of applications, and they will no doubt be increasingly used to do more than to record birthday parties.

The future therefore belongs to communications services that can provide parallel information tracks, and pack them as tightly

as possible for the brain to absorb.[22]

Multi-channel communication should lead to new forms of communications language. More pictures, symbols, and video clips will be part of what we now call the written language, because this can speed up the absorption process, and combines the abstraction of written language with the speed of visual image for other types of messages.

Of course, Chinese and Japanese have been using some symbols for a very long time. But their ideograms are frightfully hard to read and write. So it is an inefficient system. But the basic idea may be sound. And new information technology makes it possible to simplify the use of symbols considerably, because one can input them much easier.

At the same time, the demands of automated information processing will impose certain standardization of format and even syntax and style to many forms of communication, in order to make it intelligible to machines. Authors may ignore these demands, but only at their peril, because their writings may never make it to most readers. It is as if they wrote in Hungarian.

Thus, the written language itself is likely to be changing with technology, and with it how we speak, think, and interact.

---

[22] Take voice telephony. It is very inefficient in terms of information flow per time unit; which is why people can be on the phone while cooking, driving, and daydreaming. Fax is much faster than voice. Some deli sandwich places now accept orders by fax, because it is speedier and more accurate to transmit the order information in this way than by voice phone. Also, you can hire someone who speaks little English to fill the order.

6. **Information Screening**

This leaves the last and major strategy for dealing with information overload: **create screening mechanisms.** Here, the alternatives are

(a) Screening professionals, such as editors.

(b) Intra-organization screens, such as secretaries and staff. As President Reagan proved, one can boil down any issue under the sun onto one index card. It helps, of course, to have three million people working for you.

(c) Use economics as a screen, for example, by imposing an access charge on senders. Why is our time a free good for anyone who wants to access our mailbox or telephone receiver? Why should we be at the mercy of someone else's ability to get to the top of an in-box?[23] Let them pay for access. In the upper reaches of power and prestige, access was always paid for directly or indirectly. In Washington that has contributed to the emergence of the Congressional industry. The future may see a spreading of markets for access time for regular folks, too. For example, we could have personalized 900 telephone service that forces anybody calling us without knowing a personal access code to pay a credit into our telephone account.

(d) Most important, however, is an automatization of the information screening process. This is arguably the key

---

[23] In attempts to get priority of attention, the senders of information tend to employ transmission means that imply great importance—express mail; fax; influential lobbyists. But in time those efforts cancel each other out, though creating in the process a more wastefully expensive communications process.

technological challenge for the information sector. The super pipes requires the super screen, and the personal network requires the personal screen. We need technologies to help us get only information we wánt or need. The main value added may become the information subtracted.

One example for a very simple screening mechanism is a personally customized newsletter, which has only information that one individual is really interested in. For me, for example, it would include only items on Columbia College football victories and, until recently, the Albany night life. Of course, this makes for a very short newsletter. So the screen works. As everyone who used a data base can tell, the difficult part is how to suppress repetitive or unimportant information. That is, one needs a screening by quality and incremental value. Expert systems and artificial intelligence applications will be useful here, but one should not hold one's breath for their arrival. Screening is in its infancy. Right now, no computer in the world can summarize a text. No computer in the world, at any price, can write one of those four line plot capsules for TV Guide. Furthermore, meaningful information screening is highly personal, because even sensational news is an unimportant item to a person who has heard it 5 minutes earlier.

Thus, information screening requires a lot of brute force matching of the new information with the already existing

information base.[24] And that requires personal supercomputer capacity of huge strength, storage and mobility. Transacting with them will require large bandwidth. Today, everyone in the telecommunications industry is worried whether all these fiber lines will be filled, and would pay for themselves. Those people worry about the wrong thing. Of course the pipes will be filled, but only if there is a decent screen available. The problem is not the addition of information, the problem is the subtraction of unnecessary information. If one can screen the informational garbage out at the output stage, one will get the garbage in at the input stage, which means traffic for networks. Therefore, the golden rule for communications networks is: **garbage out, garbage in.**

The main point of the screening discussion is that openness in networks and in information flows, those aims Ithiel Pool stood for, are stymied without an opening of the last bottleneck. If one cannot open those 20 inches, by technology, education, and changes in the way one communicates, the rest of the system will back up like a sewer pipe. Electrical engineers speak of "impedance" as a measure for a generalized resistance of a circuit. If the impedances of parts of a system do not match each other, energy transfer is inefficient. By analogy, we may speak of

---

[24] Such a screen need not mean a total screen-out of unconventional or unexpected types of information, leading to a narrowing of users' intellectual perspective. A screen could include some randomness in the access of entirely new types of information, or a connection to the screen of other people whose judgment one trusts, or to designated editors.

"information impedance," a non-matching of the impediments to information flows.

For transmission carriers, the creation of a screen is important, because its existence leads to increased information traffic and hence business. They should support the development and creation of such screens. But, as analyzed earlier, the network environment is becoming pluralistic rather than centralized, and there will be an underinvestment by network operators in new network evolution. The reason, analyzed in our earlier discussion of network dynamics, is that the first entrant in creating a critical mass for a new network service takes risks, while the benefits are reaped by others. Even where the screens would be placed inside a network, their existence would benefit all other interconnected networks by making their information flows more manageable. In consequence, there will be underinvestment in network-based screens.

VI. Outlook

We have come to several conclusions:

1. The success of the common network in creating broad-based communications undermines the foundation of its own exclusivity.
2. The exercise of freedom of association in networks undermines the exercise of freedom of speech.
3. The advances in communication diversity undermine individual and organizational ability to handle information.

These conclusions point to bottleneck instability, conflict, and new forms of restrictions, all within an environment rich with communications. What are the implications for policy? In the past decade policy was correctly focused on creating openness by reducing barriers and permitting entry. Now, with fragmentation of the network environment proceeding apace, the primary issue is to create points and rules for **integration** that permit the continued interoperability of a "network of networks." These issues are more difficult to deal with than the policy questions of the past. We now have to construct a post-deregulatory environment, a task much harder than the initial revolution of liberalization. And, one must assure the freedom of speech to be maintained in a federated and pluralistic network, just as it has been on the unified network.

For society to deal with these issues requires foresight, knowledge, and conceptual thinking. And it is here, especially, that we all miss Ithiel de Sola Pool, who has spoken so long and so well about the technologies of freedom, and who keeps speaking to us, challenging us to remove the bottlenecks in human communications, old as well as new ones.