CONCLUDING SUMMARY: WHERE DO WE GO FROM HERE?

MR. EDSON: As the "summary session" implies, we are here to have a parting glimpse of where we have been during the last few days, and, as the program states, perhaps a few comments on where we might appropriately go. I would not continue with that one with a three-meter poll. (Laughter). But the nice thing about a summary session is that it is sort of like celebrating the 4th of July: You save your biggest, finest skyrockets for the final five minutes of the skywork display, light the fuse and stand back. It is always exciting. We have, I think, seen the results of a lot of hard work on the part of our panelists here, and I have asked them to take this opportunity to present any additional thoughts that they have concerning this meeting. I would expand the title from "Where do we go from here" to perhaps "What is missing? What should we include in any future meetings?"

With that, we are going to just sit down, relax and let it all hang out. So, do not expect anybody to stand up here and give a formal talk. We will start off with Jon Leverenz, who had the economics panel.

MR. JON M. LEVERENZ: Thank you, Dean. Good afternoon. It certainly is nice to surface. I came up on Monday, and now I am surfacing again after the Ray Boyle show, and that is quite an act to follow, but I do have a few comments to make. First of all, in our economics panel we attempted to put together sort of overall costs on a cross-section of various groups concerned with automated production systems. As you remember, we had the large federal system, we had a specialized university, and then I talked about a very small segment, and maybe I did not make that clear, very small segment of the commercial map making industry in the United States.

Now, unfortunately, we did not have a representative from the group of map makers or maybe a number of representatives from the group of map makers concerned primarily with, say, public utility maps, aerial surveys, the map companies and public utilities that usually produce large scale one-color highly specialized maps for maybe in-house use or for a relatively small special interest market. That is unfortunate, and someone brought that point up. I hope for the next panel or even before, we can make sure that a representative from that group is on the panel and, certainly any other groups that are interested in automation and automated systems.

When we presented our talks, each of us, I think as I looked at it, did try to define the particular needs of each group or agency or firm. I think that was very important. That is, 1.) the purpose must be well defined before we can begin to develop methods of a achieving that purpose as far as automated cartography is concerned, 2.) defining the detail we need in the data bank and the detail of output and the particular needs of the market that we service, 3.) we also talked about special factors which affect the choice of equipment. Then we did, finally, furnish some documentation of the dollars necessary to implement this system. I feel that as far as I can recall, Dean, that this was one of the first attempts to really try

to get out some real good costs, overall costs, on what it would take to get a good automated system going. Because it was the first, it certainly was not complete. But we did attempt to address the dollars and cents aspect.

I think that in the future we need figures, detailed figures, much more detailed figures so that we can better evaluate methodology, equipment, the efficiency, etc. I think that these are some of the things that we should be looking at and be cognizant of as we spend these large sums of money in experimental work. We should be gathering much more detailed information on developmental costs, not just running costs and equipment costs, but how much does it take to set up a file system? What is the structure of the file system? How much does it cost to develop? And determine these costs from the very beginning of the automated adventure, not try to file back through a bunch of figures and hours and contracts, and come up with something, but use some type of structure to keep track of costs from the beginning.

The second thing I think is that we ought to look a little bit more carefully and define overhead costs. What is the real downtime on existing systems in the federal government or in other activities. How much maintenance is there? That gets into manpower and downtime again. Also, I might mention, what are the failure rates? We very seldom hear anything on the failures.

The third thing, I think this would be good for establishing credibility within the cartographic community, and that is that the federal agencies, because, of course, they are spending most of the money, should really set the pace for more explicit, detailed, yet simple total costing on developing structures, file systems, as well as the other things that they seem to have done like digitizing times and equipment costs, et cetera. I think there is a real need for this.

The fourth thing is, I think, we should begin to place a dollar figure on that old heavily weighted reason for going into automation given by many military agencies, and that is speed, or "our mission." There should be something placed upon that, some dollar figure or some time figure or what it means, so that after the system is initiated, we can go back and evaluate it more effectively as far as the cost-benefit is concerned.

Fifth. I think we have to get some real good feedback now that we have operational systems going from these users of production line systems concerning the real output of these systems so that we can evaluate that against what the suppliers have stated the maximum output or minimum output is, so we therefore can make better evaluations about the systems that we install. I think it is very important. And, then, finally -- I mentioned this once before -- I think we should have a session sometime on failures. With that, I will conclude my remarks.

MR. EDSON: Thank you, Jon. I am sure that in the future we will be seeing a lot of well bitten bullets. The economic impact is here now, and the whole usefulness has to be justified in terms of dollars. The bottom line is important. In fact, this is the first time that I can recall ever publicly trying to stir up that hornet's nest.

Next, Waldo Tobler, who is the chairman of our panel on display requirement. Waldo?

MR. WALDO TOBLER: Thank you, Dean. I do not have any rockets, because I did not know I was going to be on this panel until the final program. But I did make out a wish list, and made some speculations on what we might talk about at AUTO CARTO IV. I have no official connection with AUTO CARTO IV, and I do not even know if there is one planned, but first of all let me give you some of my biases. I am in a university and I do teaching and research. That means I have no operational responsibilities, essentially no budget, and the market I service is students. Therefore, I do not have any systems, and I hope I never do have a system.

In terms of research, again, I am not very interested--this is a strong bias of my own--in large amounts of data, but I am interested in very carefully selected data, and I do fairly refined analyses of this data. I do not think Newton had very much data when he came up with the law of gravity. Well, I do not expect to come up with the law of gravity or anything like it, but it is a bias in that direction. On the other hand, the operators, the people with operational responsibilities typically use large amounts of data, but do very simple analyses. Somebody commented, for example, one could use a computer to change map scale. Well, as far as I am concerned, changing scale is a simple multiplication, and that is something one learns in the third grade. I cannot get too excited about things like that.

In terms of hardware, hardware tends to get obsolete very quickly. In some ways I think I am fortunate in not being able to buy any, because then I would wish I had the best all the time, and I never will. A $2 Etch-A-Sketch works very well to teach computer graphics. (Laughter.) For $20 you can get a little more elaborate toy. But, there are some devices I would like to see.

One longstanding one that I would like to see is the wrist watch latitude-longitude indicator. I do not know if I will live to see that or not, but it will probably come sooner than most people think. The other one that I have thought about, and have actually tried to write a program for, although the device does not yet exist, is the pocket calculator with a little LED screen, two inches by

two inches, that can do contour maps. The reason for this is, as a scientist I often get not very much data, but I want to look at it quickly and think about it, sort of the back of the envelope kind of calculation. A little pocket calculator with which I could do contour maps very quickly would be a very useful thing to have.

The mass storage devices and the high speed processors, and so on, are all right, for operations, but I am not so sure that they are going to be so interesting to me for research. In terms of firmware I have a few other items on my wish list. I would like to see a hard device, for example, that one could plug in the back of a Tectronix 4051 that does geometry and topology. Example, line intersection and distance between points, area of polygons, intersections of polygons, polygon overlays, polygon shading and that sort of thing would probably be very useful. I think there are some possibilities here, and I think Ray's panel did not address that very directly. Another wish list would be in the area of data. Each of us, I think, could come up with some. One set of data would seem to me to be very useful in this field is some statistics on what is going on. For example, I would guess, and I would be willing to bet on this, that the number of nodes in a city is a log-linear function of the population with an R square of at least .8, probably higher than that. It would seem very useful for the profession to have a number of rules of thumb like that on which we could make judgments of what costs and so on are going to be.

Another thing in the area of data. As you all know, the most rapidly growing sector of the computer industry is word processing. That is because people use words, not data and not numbers. I expect that within a very short time most electrical typewriters will be equipped with a plug where you can pull the phone jack out that put the typewriter plug in, and you will essentially have a termainal at home. Technologically that seems very simple. You would dial up by typing the numbers on the keyboard. This means that virtually everybody in the country would have a computer terminal. Now, the question then becomes, what is the data base--for example, in the last session we heard that some California county had put the Assessor's records in the computer. Now, that is public information. Am I going to be able to tap into that data directly from home, from my office, find out what my neighbor's house is assessed at? Presumably, I can do that now. But could I do it from a terminal at home? Again, I do not think that the hardware people address that question directly. That is partly not an equipment question but it is really a policy question, which is going to become more severe I think.

Finally, in the area of software. One can easily think of such things as device-independent software, software certification and

publication, software standards and so on. It is also pretty obvious that most of the software is oriented at doing the things that we have been doing in the past, not taking advantage, really, of the technology to do new kinds of things. For example, I notice the people who have the Assessor's files in the system do not normally make contour maps of assessed value. I have seen two such maps done by computer. Nor, for example, are the people who are putting population density in the computer doing the second derivative population density maps, probably because it has only rarely been done before.

Another area of particular interest to me personally was covered in one of the panels on Monday; Harry Andrews gave the work they are doing with picture processing. This all operates essentially on pixels, picture elements. These are little square polygons, each of which has four neighbors. Now, I would like to generalize some of those operations to census data. As you know, the census gives us a very blurred picture of the United States. That is because they have these little polygons in which they aggregate data. I would like to sharpen up those pictures in the way that Harry sharpens up conventional pictures.

The easiest way I find to explain the difficulty is to notice that all of those operations essentially work on neighborhood relations, and every picture element has four neighbors. It is just like a chess board. Now, you can image trying to invent a game of chess to be played on the county board of Virginia Counties, for example. Take counties in an Eastern state, any one. Imagine that it is a chess board; color it red and black. Now, devise a set of rules to play chess on that board. You have to generalize the conventional rules of chess, which are very simple rules. Playing chess is difficult, but the rules themselves are simple. A somewhat comparable operation would be to feed two pictures into the computer, taking two time slices, that is, you could take two pictures of a chess game at two different times during the play. Now compute the rules of chess from the pictures.

Well, I would like to compute the rules of land use change from two land use maps at different times by feeding them into a computer. I have actually had a student (S. Guptill) write a thesis on this soft of thing. That is what I meant by more a refined type of analysis. Again, the natural way for a scientist to compare two maps is to do cross-spectral analysis. Well, how do you do cross-spectral analysis when the data are given in the form of polygonal census data?

Also very interesting to me is the work on things like the algebra of qualitative data. The algebra of non-numerical data. These

are the kinds of things I find interesting. But, thinking a little further ahead in terms of software and what one might do at AUTO CARTO IV, I asked myself, do there exist problems that I do not know how to solve today? Are there any problems I can think of that I couldn't figure out some way to solve? I find it very difficult to think of such a problem. The question is really trickier than it seems, because we really have to talk about problems which are solvable, and you have to define what you mean by "solvable" in a rather extensive way. But this leads directly into what computer scientists call the study of algorithms and it relates to the question of algorithmic complexity. For example, just as a simple study of algorithms, everybody knows you can convert from polygons to grids and back. Right? Well, let us do it again. Let us go back from the new set of polygons to grids and back and back and back and back and back and back--and what happens in the limit? I do not think that kind of question has been studied very much. It really relates to the inverse of some of the algorithms that we have worked with. For example, there are street address to state plane coordinate programs. I do not know of anybody who has a state plane coordinate to street address program. There are all sorts of things one can imagine along these lines. But that is still a very simple problem. Polygon overlays is often considered a complicated problem in this field. Now, I think--I am not sure about this--but it seems to me that the polygon overlay problem goes up as the square of the number of polygons. You have N polygons times M polygons, and if you increase those two numbers you simply multiply the problem. That suggests that it is a polynomial problem. Now, it is well known in the computer field that polynomial problems are solvable. They are not like what is known as exponential problems, which are essentially impossible to compute. So, it seems to me you have been talking mostly about solvable problems here. But there are interesting geographical problems which do fall in other classes. This relates to theoretical work on algorithms. For example, it is well known that the best possible algorithm to do sorting will take N log, N time. There is no possible algorithm that can beat that in time, and if you have an algorithm that does it in that time you have the best possible time algorithm.

There are some problems that are even more difficult. These are generally called NP complete problems, and, somewhat crudely stated, you can prove that there are no efficient ways of solving them. One of the ones that comes up in cartography is to program a plotter to minimize the plotter movement. That is essentially the same as the traveling salesman problem. It is now believed that there is no efficient way of doing it. The only way of finding the optimal path is testing every possible path. I think we can ask if several people have polygon overlay problems what is the

theoretical minimum time that that should take, and how close do these algorithms come to it? It is a little more complex, because you have to worry about storage trade-offs, too, and so on. My point here is that there is a whole class of things that we did not talk about at this conference, and they might be topics for next year's conference. Thank you.

MR. EDSON: Thank you, Waldo. I did recall that, I believe, it was at AUTO CARTO I that Waldo appeared in fact with a $20 Etch-A-Sketch and appeared on the hardware panel. (Laughter.) Duane Marble, as the moderator of our software panel, I am sure has a few concluding remarks. Duane?

DR. DUANE F. MARBLE: Thank you, Dean. The software area is a rapidly developing one. Many of the comments that I would like to have made have already been made by Waldo. So, I will content myself with making some general statements in this area.

There is a lot of software out there, and a lot of it is very poor, and we really do not know very much about the operations we are undertaking. Sit down and ask yourself, if you were to undertake a sequence of cartographic operations and you wanted to write computer code to do these, how would you find out the most efficient way to do this? At the present time there is no good answer to this question. This leads us into the area of what Waldo was talking about -- that is, algorithms. We do not know enough about the algorithms or, to use an alternate word, the procedures that are necessary for development of cartographic software. I talked to you earlier about the IGU/USGS inventory of computer software. These programs in and of themselves may perhaps be useful to individuals, but the most useful thing in them is the ideas, the algorithms that underlie them, the ways of doing things. In many cases, the statement that, "I don't really want your lousy code, what I want to do is steal your ideas," is an appropriate one. But at the presnt time we do not have a library of cartographic or spatial data handling algorithms. I think this is something that we need to undertake as a critical research area. This involves a lot of things, some of which we find it difficult to do because most of us are basically cartographers or geographers and have not really been trained to think in computer science terms. This is a problem that is going to have to be overcome. We need to pay a great deal more attention to algorithms and we must do it in a formal

Sense. We have to start worrying about some of the algorithmic problems that Waldo discussed.

Two of the sessions in the software day were oriented toward problem areas. One dealt with raster processing of cartographic data. The other dealt with the handling of large volumes of spatial data using data base management system technoloty. Both of these areas need a great deal more examination. I think the raster approach is an example of one of the things Waldo was talking about, and that is trying to find new ways of thinking about things. I do not think we should constrain ourselves to trying to automate existing techniques. I guess for that reason I find myself a little uneasy with the term "computer-assisted" cartography that was used in the opening day's session, because it has the implication, at least in my own mind, of using the computer to do just the things we are already doing. The computer is a new tool. There are new things that we can do with it, and we must search these out.

The comments that were made about costs are also important. In the software area we have the ability to do a great deal to influence the economic viability of these operations. But at the present time we do not know anything about the cost functions. The IGU Commission tried to do a study on comparative digitizing operations. So did the Corps of Engineers, the Forest Service, and for all I know, some other people have as well. Most of these studies have not produced what we really need in an operational sense. I sit here, for example, with two maps. I can say this is a simple map, and you will say, "Yes, that is a simple map, it doesn't really have very much on it." I say this other one is a complex map. You'll say, "Yes, that is a complex map because it is the Grand Canyon topographic plate," or something like that. Obviously, there is a difference between these two maps and the effort that is going to be needed to encode them. What an operational manager needs is a digitzing cost function that will enable him to take a given set of map data, either in archival form, which we have been talking about implicitly, or direct digital capture form, which is something else, and say how long is it going to take to do this, to create the data base, and how much is it going to cost? You cannot very well go to a board of supervisors or the Federal Bureau of the Budget and say, "I think the project is going to cost somewhere between 60 and $600 million, and we will let you know exactly when it's done." The way we should do this is to find out what are the appropriate measures of information on the map, which of the measures are im-

portant factors in the cost function, how do the cost functions differ for the different digitizing technologies -- the table digitizers, scanners, automatic line followers -- each one of which will have a different cost functi One should be able to go through, take the maps, take the spatial data set, devise a sampling scheme to measure the parameters, insert the parameters into the cost functions, and come out with an estimate as well as a variance measure. Until we can do this, it is going to be a very risky operation building a large spatial base.

The same thing occurs when we start getting into operational software. If we do not know about algorithm efficiency, if we do not know about efficient data organization and management, then we cannot translate these into dollars and cents and days and weeks and years of time. This means we are not going to have economically viable operations. The development of this economic information is very difficult at the present time. We have to do our homework in other areas first, not only in the hardware characteristics but in how we measure spatial data and how we manipulate it and store it.

This brings me to two concluding topics that I would like to pose as questions. First, there seems to me to be a need for a substantial amount of cartographic research, applied research or day-to-day engineering problems. But there is a great deal of work to be done out there. Some of these things will be done by operational agencies under the pressure of their line responsibilities. Other portions of it will not get done that way. I worked in the urban transportation field for many years. There was a constant problem of who was to fund the needed research. The things that agencies like the Department of Transportation felt were blue-sky basic research, in the eyes of many academic researchers appeared to be quite applied.

Many agencies are unable to fund things that cannot be directly related to their day-to-day operations. But someone, at some time, is going to have to start worrying about integrated research programs in cartography and the provision of financial and manpower resources for carrying them out. At the present time I think, for example, the National Science Foundation would be hard pressed to fund things in the software area. They have a division that deals with computer science. We have no credibility in that area. We are not computer scientists. They have a program that deals with geography and regional science which funds no work in cartography. The engineering division, I do not really know about, but I doubt their in-

terest as well. There is a very real problem with funding and organizing basic research. I think the people working in cartography and geographic information systems are going to have to worry about this.

The problem is not only that of funding; there are questions of manpower -- manpower for research and manpower for development work, and manpower for carrying out day-to-day operational responsibilities. I think that, again, we must ask ourselves what are we doing in terms of training programs? Again and again, during the meeting and other times, people come up to me and ask do I know anyone with thus-and-such a set of characteristics. Most of the time I say, "Yes, but they are already gainfully employed." They ask about students who are trained, for example, in cartography and computer science. There are very few of these that are coming out. I think we in the universities have to look at our training programs and see what we can do to restructure them to try to start solving some of these problems. We also need large scale training programs to upgrade skills of existing staff. There are a lot of people that are working in the area of computer cartographic today that came into it with no formal training. I am one of them.

The things that are learned on your own sometimes sink in deepest, but they are also very hard to come by. There is no effective program of in-service training on either the Federal, state, or local level. I feel things have to be done in this direction as well.

We have a number of things to do in the next few years, a number of problems to attack on both an operational and sort of basic research and training areas. I would certainly hope that we can meet them efficiently and not just stumble ahead into the future. With that I will pass it back to Dean. Thank you. (Applause.)

MR. EDSON: Thank you very much, Duane. Perhaps in the next couple of years before AUTO CARTO IV, we will have to consider a new term, perhaps "computer-biased" cartography or something like that would be better.

DR. MARBLE: I would like to suggest "cartography."

MR. EDSON: That sounds super to me. Next, Ray Boyle, who has held forth yesterday and part of today on hardware. I am sure Ray has some concluding remarks. Ray?

DR. RAY BOYLE: I think I did most of them this morning, and Duane has done quite a number for me where I was going to touch into the software area. It is very difficult to in fact separate software and hardware. The two go along together.

One of the things that I noticed in my travels and in talking with people about programs, I have not yet come across a program that has been put out as a job to be done that specified it had to have a certain run cost. Nobody seemed to worry about how much it costs to run their program when it came back. I think that cost-effectiveness should be in every contract for programming. Hardware, particularly, and software as well, are entirely dependent for their application on their economics. We are no longer at a stage where it is possible to just buy a toy to play with. It has to do work and it has to do it efficiently and economically.

We are starting to get the feel, if not all the information we need -- and we tried to bring some of that out in the hardware panel. I feel confident that scan digitizing can meet the economic requirements that are needed. I also feel that the smaller optical disk memory will be able to meet the necessary economics of storage and transmission of data. I would like to feel that these will be used because I believe that they are an essential part of cartography in the future.

Somebody remarked to me that it seemed to them from the meeting that the poor old vector is dead and now we have only to work in raster scan and think only in raster scan. This was no intention of mine in the hardware part. I do not think it was any intention in the software part either. However, there has been a stress in this presentation of scan methods. I think in some areas they are better; I think in some areas they are worse. Perhaps, because we are presenting new things at this meeting, there has been a slight overaccentuation of the word "scan." The vector is there, and it is good, and we have got a lot of use to come from the vector method. I think so myself for cartography, for cartographic storage. But I am prepared to be argued out of this. There are arguable matters in both directions.

The audio aspect of hardware I think is important, and I think will be used more and more. It is getting to be very low cost; it is simple and good.

When I talk about the small optical disks for memory, I am trying to avoid LANDSAT imagery, storage and ideas. I am treating those

as outside automated cartography. I think that for ordinary cartography something of the order of 10^{10} bits is good storage. As I said this does mean that you can store about a thousand topographic quad sheets on one rotating disk. Quite a nice handeable amount and good quality data that will be with sufficient redundancy checking so that there will be no problem of errors.

With regard to the question of large displays, in my travels I have certainly found that in presenting the results of geographic information systems to politicians, they demand the large display, as Dean Edson said in his final remarks. I am not certain whether it is wanted in editing work in cartography, and we did not get a clear answer in our panel meeting. The question was left in the air. Maybe that is where it should be until we see the costs of these larger screens and how we can use them effectively.

I was extremely happy, and I would very much like to thank all the members on the hardware panel. We did explain to them how we wanted their ideas put over. They were going to talk to cartographers, cartographers who knew their subject well, but would not understand jargon in some of these areas. I do compliment them on doing an extraordinarily good job. There was never any question of my going to sleep in any of the discussions. I was enthralled, although these were things I had heard before. I think that the written account will be a very valuable document for people outside this room as well as to you, because I am certain some of the things you will read again to take them in. Ideas came thick and fast. They were all good. I shall be very happy to read the report when it comes through, and I know I will be meeting a lot of people who will be asking for it. I hope that many copies will be made available, Dean; we ought not to have such a limited number that it goes out of print one month after the first issue.

I still go back, I think, to my original comments, that I am not certain that this is not going to be the peak of hardware developments. I think that most of the things I want to be answered are being answered and are being answered by engineering hard work, not by a new vital breakthrough. We did not cover every subject, but we covered a lot of the subjects; we had to be selective.

I would also like to repeat the comments that I made, the summary remarks this morning, that I do believe that there has to be a greater effort at getting good education to cartographers, updating the thinking in these areas. It is a new type of cartography. You do not have to learn programming, but you have to learn a new way of thinking. I hope that the universities can perform this teaching job for you as on-line working people, and that departments will support this development and implementation of in-service training.

I also hope that the government will support the idea of service bureaus. There are a lot of people who want to use these facilities. They are not going to be able to run them 24 hours a day or even one minute a day. But they still want to use them occasionally. So, please, do, as government administrators, please think of this need from outside.

I am very happy with the state of hardware as it is, both with the manufacturers who are showing here and things that I have heard have been done. I hope that you also feel the same way. That is all I want to say. (Applause).

MR. EDSON: Thank you, Ray. I am very pleased that you altered your position slightly from your opening remarks, because I feel I am almost looking forward now to the next couple of years, because I am almost positive we will see new inventions, new thinking in the hardware front. It just cannot help but be that way.

Finally, the panel on data edit, headed by Harvard Holmes, the cleaner of the non-clean data.

Harvard, what can you tell us?

MR. HARVARD HOLMES: I guess I will start by answering Ray. Our editors would like a large screen, as large as possible. We also have to show that data to politicians. I think we have to realize that politicians do make decisions, and there is a lot of worth in showing the data to politicians. I guess as a computer scientist and not a cartographer at all, my real interest in cartography is as an information transfer mechanism. So, my real interest in looking toward the future is in tapping some of the very large data bases that have been collected and using cartography and mapping to convey that information to the analysts, the planners, the politicians, the people who need to understand what that data says.

I guess I, too, would be very encouraged by the current state of affairs in cartography. I sort of get the feeling that for my needs, pretty much everything is there. I can get it. I can get the hardware, I can get the software for the cartography aspects of what I am doing. What I cannot get is the mass storage system yet. I have hope for those. What I have a little less hope, at least for the present, is for the software, to access these map data bases. It is still not clear to me if one has ten billion characters of census data exactly how a naive user can sit down at a terminal and discover what

is in those ten billion characters of data, and how to get out the 500 characters that he would like to have and get those characters out in ten seconds or so, and then in another 15 seconds have a map put on his graphic terminal so he can begin to understand what that means. I think the challenge for the next few years, at least for me, lies there.

MR. EDSON: Thank you, Harvard.

(The meeting was then formally concluded, with remarks by Mr. Edson and Mr. Chamberlain.)