LINK AND PAIR (CDC-3600 PROGRAMMES FOR THE ASSOCIATION OF SPARK IMAGES INTO TRACKS)

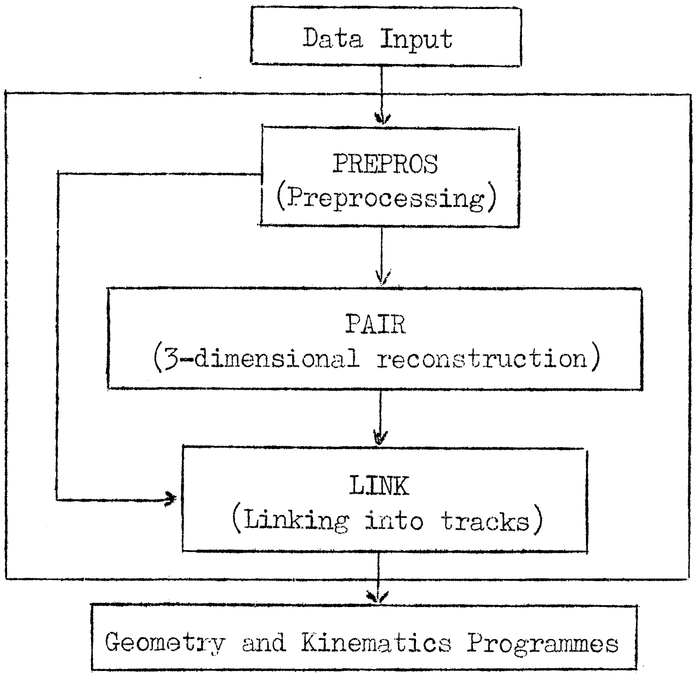
R.F. CLARK

Argonne National Laboratory, Argonne

AIRWICK is a spark chamber data processing system being written at Argonne for the CDC-3600. It is able to handle data originating from photographs, vidicon systems and sonic or wire chambers and will furnish a set input into geometry and kinematics programmes. There are three sections, each having increased generality in the system. In order, they are:

1. PREPROS - Preprocessing
2. PAIR - Three-dimensional reconstruction
3. LINK - Linking of sparks into tracks

AIRWICK SYSTEM



We believe that an important contribution to date has been the establishment of this processing order rather than the normally conceived one of PREPROS, LINK and then PAIR. In three-dimensions tracks which intersect will almost always intersect because of a vertex. The physical separation of sparks will be at least as great as in the projection and in most cases will be significantly greater, an effect which tends to normalize distances between sparks.

As we describe our approach, it should be remembered that dealing with a problem by computer often necessitates techniques which are not obvious and are certainly not those a human would use if faced with the same problem. Many of the ideas and much of the terminology we use have come from graph theory and from the applications of decision theory to pattern recognition. Two references are The Theory of Graphs by Claude Berge (Wiley, 1962) and Decision-Making Processes in Pattern Recognition by George S. Sebestyen (Macmillan, 1962). Ideas from these two areas have not been given sufficient consideration in attacking the problems of automatic data processing.

The approach taken in PAIR originated when we undertook the automatic processing of film taken by the Argonne Group at CERN in 1962. A typical stereo pair is shown in the photograph, Fig. 1 a and b. Using this stereo pair for an example, we will describe the basic reconstruction process.

Upon entry into PAIR we know in which gap sparks lie and can restrict our attention to one gap. We order the sparks in each view and can now visualize the situation (using the example below) in a 10 x 12 matrix, where the 10 sparks in view 1 represent the rows and the 12 sparks in view 2 represent the columns. Each element of the matrix will refer to a possible pairing P(i,j). Using the coordinates of the spark centroids for spatial reconstruction, assigning an element in the matrix a value of one if the new spark is in the chamber and a value of zero if it is not, we obtain an incidence matrix with a "block diagonal" property.

"BLOCK DIAGONAL" PROPERTY OF INCIDENCE MATRIX

$$
\begin{array}{cccccccccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
$$

i<u>th</u> row corresponds to i<u>th</u> spark in view 1

j<u>th</u> col corresponds to j<u>th</u> spark in view 2

ij<u>th</u> element of matrix corresponds to pairing P(i,j)

Since the sparks have been ordered, the incidence matrix has the property that once a one is followed by a zero in a row or column, all remaining elements in the row or column must be zero. This is due to the fact that if a spark in one view can be paired with two sparks in the other view, it can be paired with all sparks between them. We can separate the blocks and can consequently focus our attention on resolving the ambiguities in a single block.

We will use as an example, gap 13 from the photograph, where we have the following measured data and incidence matrix

INCIDENCE MATRIX FOR GAP 13

	(937.0)	(1030.5)	(1490.0)	(1789.5)
(1020.5)	1	1	0	0
(1192.5)	1	1	1	0
(1204.5)	1	1	1	0
(1653.5)	0	0	1	1
(2016.0)	0	0	0	1

We define a decision function on the pairing P(i,j) which takes into consideration:

1. The value of the ij<u>th</u> element of the incidence matrix.
2. The widths (intensity) of the sparks.
3. The order in which sparks appear in the gap.
4. The "clarity" with which we "see" each spark.
5. The position of the spark in the chamber (which will not be considered in the following discussion).
6. Weights, which are programme parameters, for the factors given in 2 through 5.

Functional values will range from 0 to 100. The larger the value, the greater the chance of an unambiguous reconstruction.

Below we show how the portion of the function pertaining to the widths of the sparks is calculated. Normalized widths are computed by dividing the spark width by the sum of the spark widths in the same block and view.

CONTRIBUTION OF WIDTHS

$W(i,j) = A(1 - |w_1 - w_2|)$

$A = 60$

w_1 = normalized width of spark in View 1

w_2 = normalized width of spark in View 2

	(17)	(10)	(11)	(12)
(10)	55.7	55.7	X	X
(4)	46.1	54.5	53.3	X
(6)	49.3	57.6	56.5	X
(8)	X	X	59.8	58.6
(9)	X	X	X	59.8

We now calculate the order contribution

CONTRIBUTION OF ORDER

$$O(i,j) = B\left(1 - \frac{|i - j|}{P}\right)$$

$B = 10$

$P = \max(5,4) - 1 = 4$

	(1)	(2)	(3)	(4)
(1)	10.0	7.5	X	X
(2)	7.5	10.0	7.5	X
(3)	5.0	7.5	10.0	X
(4)	X	X	7.5	10.0
(5)	X	X	X	7.5

and the "clarity" contribution

CONTRIBUTION OF CLARITY

$$C(i,j) = K \left(1 - \frac{S(i) - 1}{R}\right) \left(1 - \frac{T(j) - 1}{R}\right)$$

$K = 30$

$R = \max(5,4) = 5$

$S(i)$ = number of possible pairings with spark i

$T(j)$ = number of possible pairings with spark j

	(3)	(3)	(3)	(2)
(2)	14.4	14.4	X	X
(3)	10.8	10.8	10.8	X
(3)	10.8	10.8	10.8	X
(2)	X	X	14.4	19.2
(1)	X	X	X	24.0

We then compute the values of the decision function and are ready to begin the selection of our pairings

VALUES OF PAIR DECISION FUNCTION

$$D(i,j) = W(i,j) + O(i,j) + C(i,j)$$

80.1	77.6	X	X
64.4	75.3	71.6	X
65.1	75.9	77.3	X
X	X	81.7	87.8
X	X	X	91.3

Since the maximal value of the decision function does not exceed all other values by 10 per cent, we calculate a "comatrix" whose values are given below:

COMATRIX OF D(i,j) VALUES

$$A(i,j) = \sum_{n} D(n,j) + \sum_{m} D(i,m) - 2D(i,j)$$

207.1	231.3	X	X
292.1	289.5	298.7	X
297.7	295.3	294.3	X
X	X	236.7	173.0
X	X	X	87.8

These values give a measure of the degree to which all <u>other</u> sparks in View 1 could pair with spark j <u>and</u> all <u>other</u> sparks in View 2 could pair with spark i. Since the minimal value in this matrix A(5,4) is 10 per cent less than all other values, we choose as our first pairing:

(Spark 5 in View 1) ⟷ (Spark 4 in View 2)

These sparks are removed from further consideration by deleting row 5 and column 4 in the original matrix. We therefore obtain:

REMAINING D(i,j) VALUES

80.1	77.6	X
64.4	75.3	71.6
65.1	75.9	77.3
X	X	81.7

RECALCULATED A(i,j) VALUES

207.1	231.3	X
292.1	289.5	29.87
297.7	295.3	294.3
X	X	148.9

Again our 10 per cent criterion makes no decision and we recalculate the comatrix. As before, we are able to make a choice and we choose as our second pairing:

(Spark 4 in View 1) ⟷ (Spark 3 in View 2)

Our processing is now reduced to:

REMAINING D(i,j) VALUES		RECALCULATED A(i,j) VALUES	
80.1	77.6	207.1	231.3
64.4	75.3	220.5	217.9
65.1	75.9	220.4	218.0

No pairing decision can be made from the first matrix or its comatrix. We therefore return to the matrix of D(i,j) values. Since all values within 10 per cent of the maximal value do not lie in the same row or column, we select the maximal value and our third choice is the pairing:

(Spark 1 in View 1) ⟷ (Spark 1 in View 2)

We now have:

REMAINING D(i,j) VALUES	RECALCULATED A(i,j) VALUES
75.3	75.9
75.9	75.3

As before, the 10 per cent criterion fails to separate a pairing in the two matrices. We return to the first one and find that all values within 10 per cent of the maximal value lie in one column, and we make a simultaneous fourth choice, pairing:

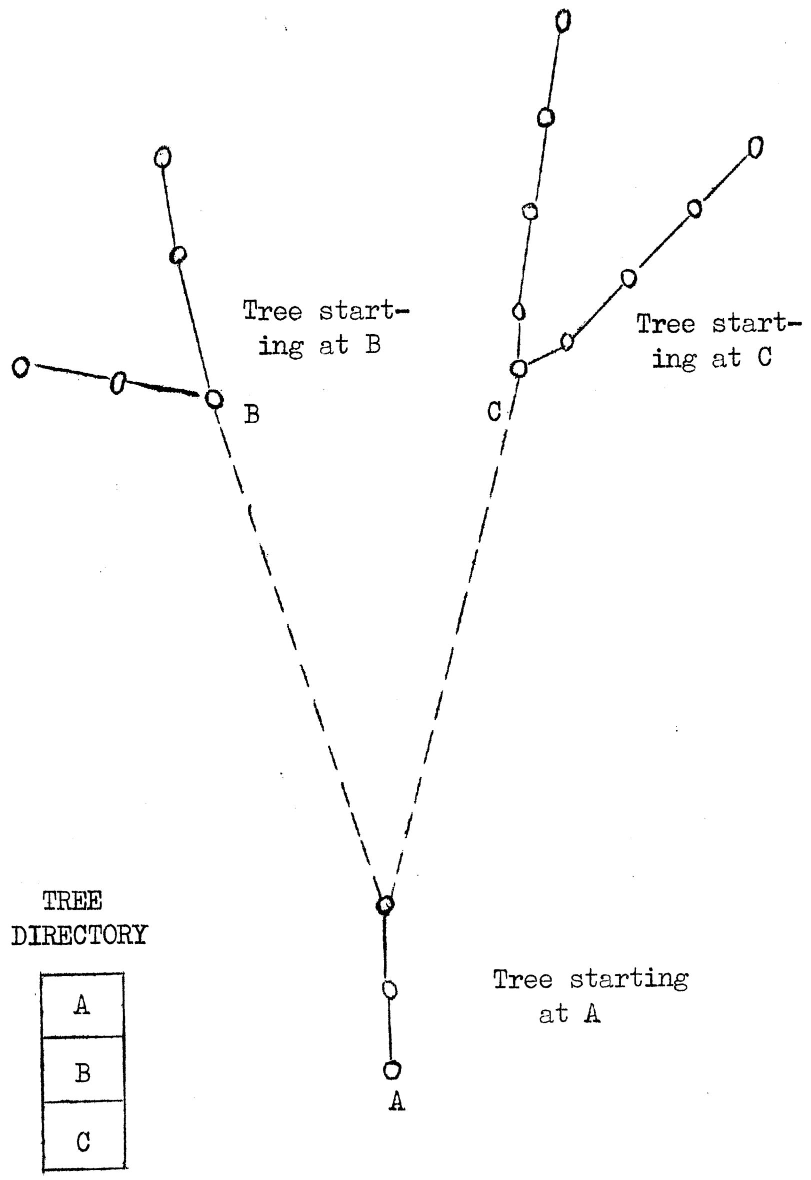
(Spark 3 in View 1) ⟷ (Spark 2 in View 2)

(Spark 2 in View 1) ⟷ (Spark 2 in View 2)

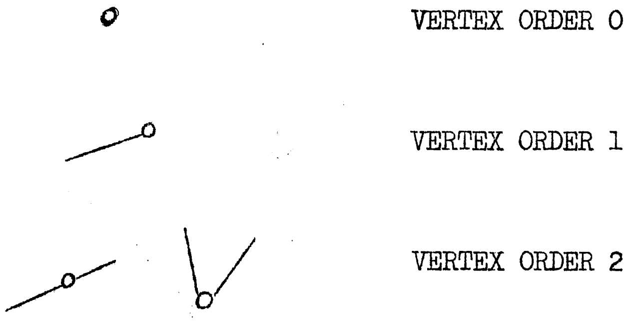
This exhausts the matrix and we have now established pairings for all the sparks in both views. If we look at the photograph, we see that these are exactly the choices a human scanner would have made.

LINK has been written to accept a random list of coordinate triples and with some knowledge of the chamber but none of the topology of the event, link these points in space into tracks. Our approach, is at present a local rather than a global one, assembling sparks into track segments and track segments in tracks.

Output from LINK is in the form of "trees" with a directory indicating the root of each tree.

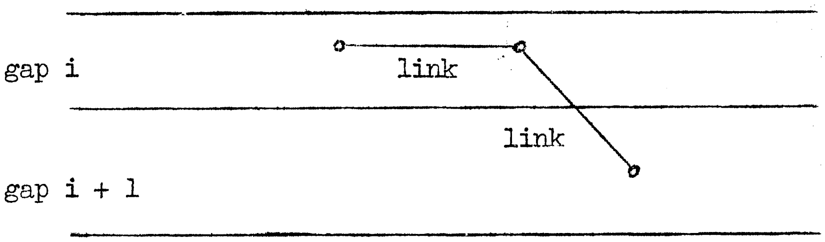


Associated with each spark will be a vertex order and links (if any) to adjoining sparks. A vertex order 0 indicates an isolated spark, 1 indicates a track end-point, 2 indicates a spark interior to a track or a vertex of a "V".



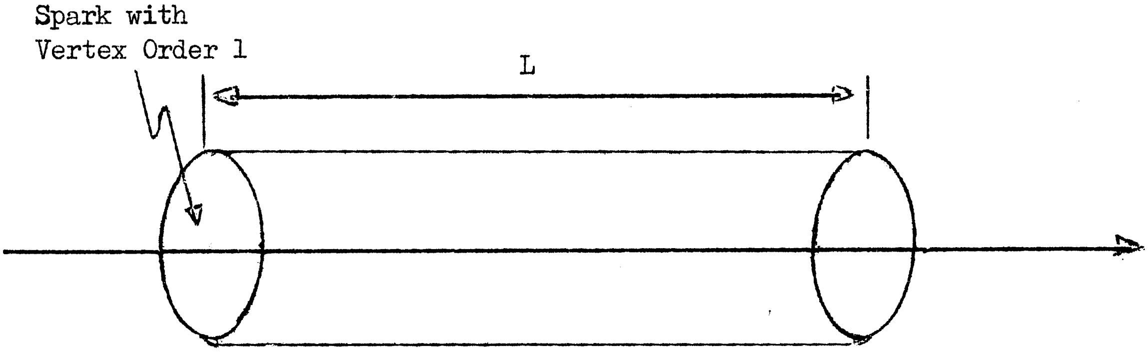
A vertex order N $\geq$ 3 indicates the vertex of an (n-1)-pronged "Y" or an N-pronged "V". This structure gives a reasonable description of the actual physical event taking place in the linking region and is amenable to certain generalized data processing techniques. The vertices which are of interest to the physicist will lie within some small sphere about the vertices of order 2 (or more) in the tree but, in general, these vertices will not coincide.

For each of ten sub-regions in the linking region, we define a linking distance W. In each sub-region, we execute a loop through gaps 1, 2, ... , N, establishing intra-gap links between sparks within gap i whenever the distance between them is less than W. An attempt is made to link sparks in gap i with sparks in gap i + 1 using a linking distance W. If this attempt fails, we try to establish links using a distance 2W. This inter-gap linking effort is then terminated regardless of the outcome.



If there can be at most one track in the linking region, a parameter can be set which forces the linking of all sparks using the criterion of minimum distance. In the more general case, when there is a possibility that two or more tracks can be present, this forcing of links cannot be done.

At the present time, we simplify matters by searching through our linked sparks and if a spark of vertex order greater than 3 is encountered, processing is abandoned. If none are encountered, another search is made for sparks of vertex order 1. If none are encountered, we again abandon processing. Extrapolation attempts are made from all sparks of vertex order 1, in which we look along a line between the spark and its predecessor and link the spark to the first spark found in a cylinder of radius M (where M is the maximum linking distance) about this line.



If no spark is found within a distance L, the extrapolation procedure is terminated. When extrapolation attempts have been made from all sparks of vertex order 1, we exit from LINK and the tree structure is transmitted to geometry and kinematics programmes for final processing.

We do not feel that this formulation of LINK is the most desirable one. As a next step, we would like to have some way of removing the restrictions on the vertex order, but we feel that a completely satisfactory solution will come only with the formulation of a global approach to the problem. A more detailed reference to the AIRWICK System is "AROMA-AIRWICK: A CHLOE/CDC-3600 System for the Automatic Identification of Spark Images and Their Association into Their Association into Tracks", AMD Technical Memorandum No. 64, Argonne National Laboratory, February 1964.

Figure caption

Fig. 1 a and b	A typical stereo pair of spark chamber film taken by the Argonne Group at CERN in 1962.

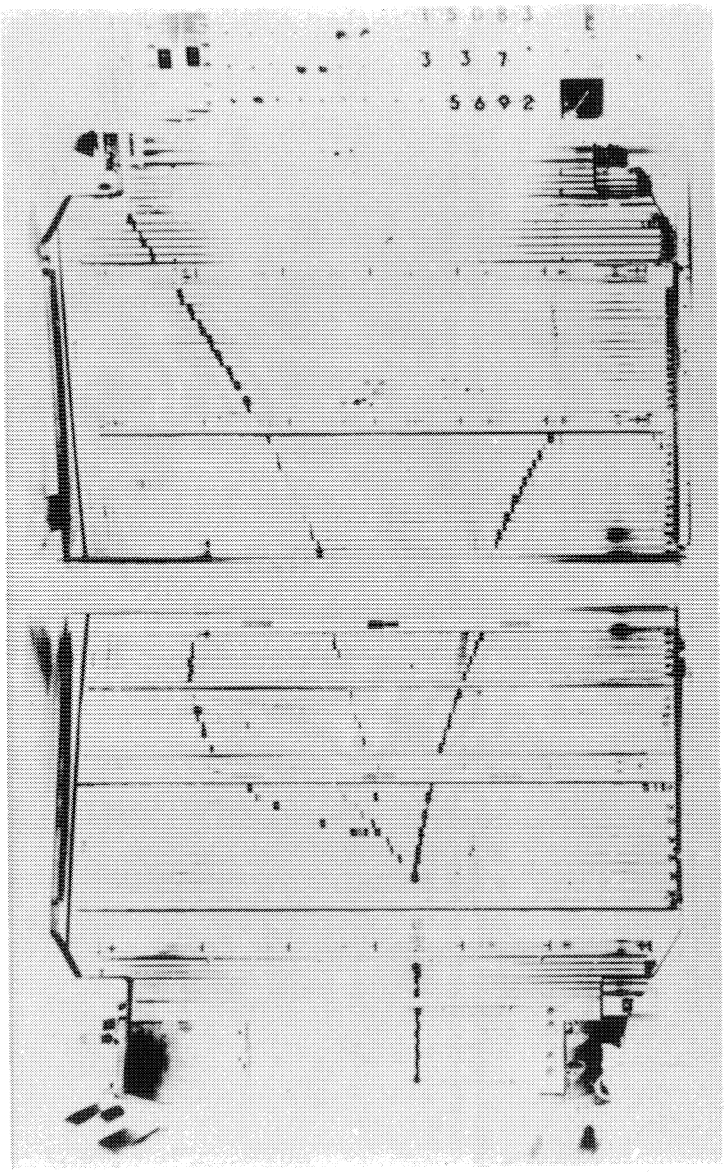


Fig. 1a

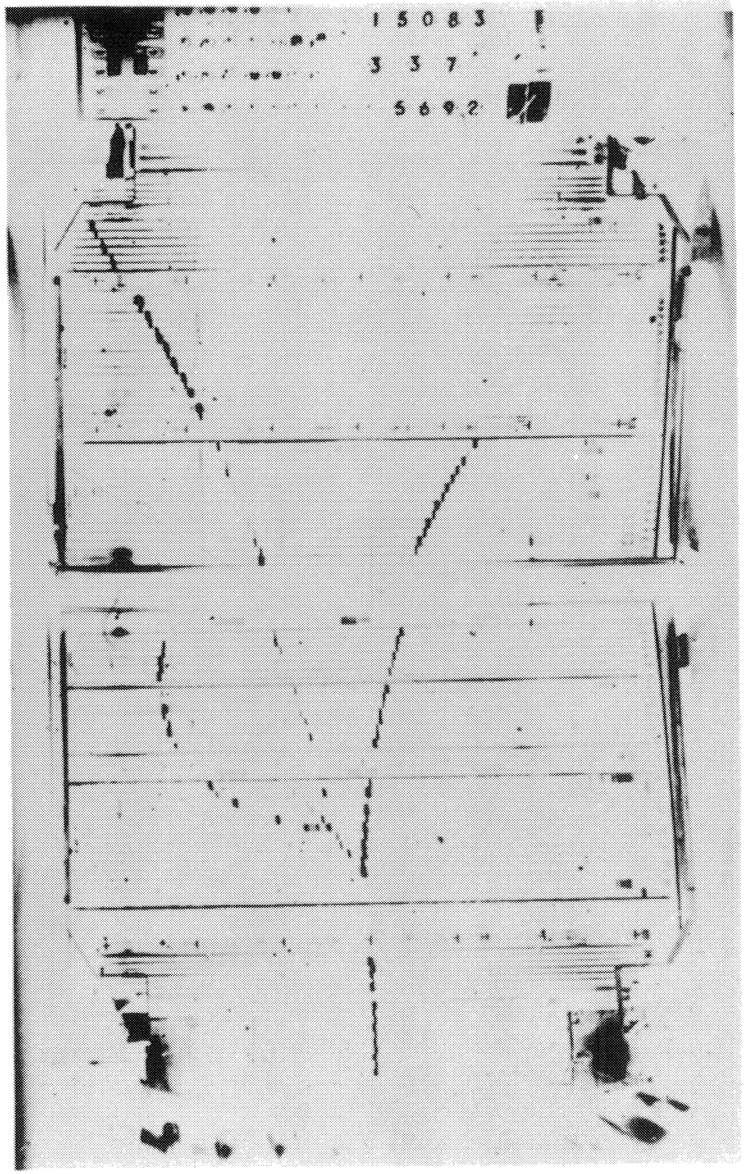


Fig. 1b

DISCUSSION

ANDERSON: Could I ask how long it takes to process one of these events ?

CLARK: For the pairing procedure approximately 1 sec, and for the linking procedure it will be in the vicinity of 3 to 5 secs. An attempt has been made to structure this thing so that the topology becomes more simple. If you have some prior knowledge of the topology, the programme may be reduced by as much as a factor of a hundred in the event of one relatively straight line, we hope.

GELERNTER: How valuable was the width information in enabling you to make pairings ?

CLARK: We experimented with the parameters to a great extent until we found the ideal weights that could be assigned to them. We would have used additional criteria had we been able to think of any that seemed to be of importance; but the one most important thing was the width contribution, the next seemed to be the confusion that the spark itself was in. If it was in an area with many other sparks where things were less clear, and if we can measure the complexity of the neighbourhood of the spark and transmit this information, it seems to be very useful. This procedure actually works quite well. The fact that you are now in three dimensions gives the following advantages: your problems of linking improve as far as computer processing time is concerned; the total computational time is less than doing two views separately and then merging them; and the separations between tracks become very great, I mean, things which appear very complex on the projections seem to disappear for the most part.

GELERNTER: Since there are such variations in intensity from gap to gap, I speak for example of robbing of energy from a spark by a high-energy particle that appears to live further down, is width information of any use in constructing a single track from gap to gap ?

CLARK: No, it doesn't seem to be. Actually, the knowledge within the gap seems to be better than that going across. You can't count on a particular spark being of the same width as one in the next gap. There seems to be very little correlation here. But within the gap certainly regardless of the stereo view whether its 90° or small angle stereo spark has a tendency to be a cylinder, and the width information is in general fairly uniform.

NOTE: On the problem of automatic scanning and track recognition of spark chamber pictures, see also reference 4) of the paper given by G.R. Macleod in Session O.

FESSEL: Do you have any trouble with tracks in a magnetic field that have spread ?

CLARK: Again, no two pictures we have processed are alike. A spiral in one of them has been nicely reconstructed in space. But I am not sure that I understand what you mean.

ROBERTS: I could comment on that, I thing what you mean is the order of displacement in the magnetic field, that is of course present in the pictures but its not very large its a few millimetres per gap and the programme doesn't have any trouble following it.