Observability and Redundancy Classification in Multicomponent Process Networks

Classification of variables and measurements in multicomponent flow networks is treated in this paper. Classification rules are derived that exploit the relationship between cutsets of the process graph and certain graphs derived from it, and the solvability of the relevant equations. These rules are incorporated in two graph-oriented algorithms for observability and redundancy classification.

Alexandros Kretsovalis,
R. S. H. Mah
Department of Chemical Engineering
Northwestern University
Evanston, IL 60201

Introduction

For reasons of cost, convenience, or technical feasibility not every variable in a process is measured. However, we may still be able to estimate the value of an unmeasured variable from other measurements through mass, energy, and component balances. Our ability to do so will depend on the structure of the process network and on the placement of measuring instruments. Mathematically speaking, if we are able to change the value of a variable without violating the conservation constraints, then the change is said to be feasible. Typically, this means that the values of some other related variables will have to be changed appropriately so that the constraints remain satisfied. For a given process network with a given set of instrument placements, if we can make a feasible change for a variable without being detected (or observed) by the instruments, then the variable is said to be unobservable. By this definition a measured variable is certainly observable, but an unmeasured variable may or may not be observable.

Clearly, it is desirable that all variables of interest to the performance of a process should be observable. For instance, all variables associated with a control loop should be observable. Similarly, all variables associated with a yield accounting should be observable. But with some applications, this requirement in itself is not sufficient. For instance, in a tank with one inflow and one outflow, we may apply feedback control to the liquid level by manipulating one or more flow rates if the level is measured. If we measure the flows but omit level measurement, the level is still observable but the control system is now feedforward rather than feedback.

If we delete the measurement associated with a given variable, and if the variable remains observable, then the measurement is said to be redundant. In a process network we require certain variables to be observable, others to be measured but not redundant, and still other measurements to be redundant. How to classify variables and measurements in a multicomponent flow network is the basic problem addressed in this paper. This problem is clearly of importance in process operation. For instance, we may wish to evaluate the impact of an instrument failure for a given process. But observability and redundancy classification also provides a rational basis for designing a performance monitoring system, trading placements and instruments of different types, costs and accuracies. The necessity for classification algorithms is made more compelling by the complexity of highly integrated processes and by the volume of data and measurements in highly instrumented and automated plants.

For single-component flow networks, networks with linear constraints, or mass-energy flow networks, efficient classification algorithms already exist (Vaclavek, 1969; Mah et al., 1976; Stanley and Mah, 1981b; Crowe et al., 1983). In this paper we extend the treatment to multicomponent flow networks, but exclude energy balances and chemical reactions from our present consideration. Previous works by Vaclavek and Loucka (1976) and Vaclavek et al. (1976 a, b), and by Romagnoli and Stephanopoulos (1980) required the composition of each stream to be either completely measured or not measured at all. The algorithm of Vaclavek et al. is graph-oriented. Nonredundancy classification rules are employed to "reduce the balance scheme" and to identify the redundant measurements. Application of this algorithm is limited by the fact that these rules constitute only sufficient conditions for nonredundancy classification. In other words, measurements that are classified as nonredundant are actually nonredundant, but some nonredundant measurements may also be erroneously classified as redundant. The algorithm of Romagnoli and Stephanopoulos is equation-oriented. Solvability of the nodal balance equations is examined and an output set assignment algorithm (Stadtherr et al., 1974)

Correspondence concerning this paper should be addressed to R. S. H. Mah.

is employed to classify measurements and variables simultaneously. Their observability classification of an unmeasured variable depends on whether it can be assigned as an output to a nodal balance equation that contains only observable variables. However, consideration of nodal balances alone may lead to incorrect classification, since an unmeasured variable not assignable as an output to a nodal balance may nonetheless be observable through balances of clusters of nodes. A more detailed critique of the above algorithms together with counter examples is presented in material deposited as Supplementary Material No. 1.

The algorithms developed in this paper do not impose any restriction on the number of mass fraction measurements in the streams. They make use of graph-theoretic concepts and proceed in a layered approach through the various graphs derived from the process graph. Variables in streams that do not lie on cycles with unmeasured mass flows are classified on the basis of graph-theoretic criteria alone. These criteria are derived by exploiting the structure of the network and the positions of the measurements. To treat the classification problem in cycles with unmeasured mass flows, cutsets of the process graph are examined in order to select the appropriate subsets of equations and determine their solvability. The observability, and redundancy, now depends on the numerical values of the measurements as well. The proposed classification algorithms have been successfully implemented on computers. This paper will focus on the development of the algorithms rather than their implementation.

Problem Formulation

The structural information in a plant can be conveniently represented by a directed graph, called a process graph (Mah et al., 1976), which we denote by F. The nodes of F correspond to the units, tanks, and junctions of the process; the arcs correspond to the process streams. The process graph contains an environment node from which the process receives its feeds and to which it supplies its products. It is always cyclic. A process network records in addition to the structure of the graph, the types of measurements as attributes of the arcs. In the networks to be considered in this study there are two types of measurements (mass flows and mass fractions) and three types of variables (mass flows, component mass flows, and mass fractions). For simplicity, we will assume that all mass fractions and mass flows are directly measured. An arc corresponding to a stream whose mass flow and mass fractions are not measured will be referred to as an unmeasured arc.

For an I-component flow network with N physical nodes and J streams, the constraints consist of mass balances:

$$\sum_{j}^{J} a_{nj} M_j = 0 \quad n = 1, \ldots, N \tag{1}$$

component mass balances:

$$\sum_{j}^{J} a_{nj} M_j x_{ji} = 0 \quad i = 1, \ldots, I; \; n = 1, \ldots, N \tag{2}$$

and normalization equations:

$$\sum_{i}^{I} x_{ji} = 1 \quad j = 1, \ldots, J \tag{3}$$

where a_{nj} is the (nj)th entry of $\underline{A}$, the reduced incidence matrix of the process graph F. The mass and component mass balances, as written above, correspond to single nodes. However, they are not the only constraints to be considered; similar balances can be written about any cluster of nodes.

Since we do not impose any restriction on the number of mass fraction measurements in the streams, the normalization equations are always active constraints. However, because they are stream constraints—in contrast to the mass and energy balances, which are nodal constraints—the treatment used in mass or mass-energy flow networks (i.e., networks constrained solely by nodal constraints) is inadequate for the networks considered in this work.

In the following treatment it is assumed that if all mass fractions are measured in a stream, they are normalized to sum up to unity. If one mass fraction measurement is missing from a stream, it is calculated using Eq. 3. The only other situation remaining corresponds to two or more unmeasured mass fractions in the same stream.

The basic theoretical framework of observability and redundancy was set out by Stanley and Mah (1981a). But the algorithms that they developed (Stanley and Mah, 1981b) apply only to single-component and energy flow networks. In the new approach the basic theoretic framework and the use of graph-theoretic concepts are retained. But rather than using a feasible unmeasurable perturbation or a perturbation subgraph to identify unobservable variables, we identify the observability through the solvability of the constraint equations. If a subset of the constraint equations is solvable with respect to the unmeasured variables occurring in it, then these variables are said to be observable. An unmeasured variable is unobservable if all subsets of equations in which it occurs are not solvable.

At this point it may be useful to recapitulate the essence of what we are trying to accomplish. The objective of this exercise is to classify variables and measurements according to their observability and redundancy. Solvability of equations is used as a tool in the classification. We are not trying to solve a given set of equations numerically, which would be a straightforward exercise for linear and bilinear equations. For most cases of interest to us there are almost always unobservable variables and nonredundant measurements present. Consequently, the complete set of equations is almost certainly not solvable. We will gain very little insight in trying to solve the complete set of equations. What we are trying to do is to look for appropriate subsets that are solvable (but without actually solving them) and use this information to infer the status of variables and measurements. It is important to emphasize the qualifier "appropriate," because we are not interested in looking at all subsets of equations, which would be numerous, but only at subsets relevant to the variables and measurements under consideration. This is normally a tedious and error-prone combinatorial problem. In our new approach the appropriate subsets are identified by exploiting the relationship between cutsets of graph G and certain graphs derived from it, where G is the underlying graph of F with the directions of arcs erased.

Biconnected components and cutsets are the most important graph-theoretic concepts employed in the new approach. They are defined in Appendix 1, where a short discussion of all concepts needed for the classification algorithms is also provided. The reader is referred to Deo (1974) and Aho et al. (1974) for a fuller treatment of graph theory. Two operations on graphs will be carried out:

1. Deletion of arcs, which is simply removal of arcs from the graph.

2. Aggregation of two adjacent nodes, with elimination of all arcs between them.

In the following treatment we shall make use of the graph G and certain graphs derived from G. G_m is the subgraph of G obtained from G by deleting all arcs with mass flow measurements. G_{mx} is the subgraph of G containing all unmeasured arcs and only unmeasured arcs. It is derived from G_m by deleting all arcs containing one or more mass fraction measurements. Thus, $G_{mx} \subset G_m \subset G$. Graph G^m is obtained from G by aggregating all nodes that lie on a cycle with unmeasured mass flows. Graph G^i ($i = 1, \ldots, I$) is obtained from G by aggregating all nodes that lie on a cycle with unmeasured mass fractions of ith component. G_i^m is the subgraph of G^m obtained by deleting all arcs with measured mass fractions of ith component. A chart relating G to its derived graphs is given in Figure 1. Illustrations of graphs G_m, G_{mx}, G^i, G^m, G_i^m may be found in Figures 2, 3, 4, and 5. Note that in these figures, the labels in the parentheses following the stream labels indicate the unmeasured variables in a stream: "M" means unmeasured mass flow and "i" ($i = 1, 2, \ldots$) means unmeasured mass fraction of component i. The same notation will be used in all subsequent figures.

Classification Principles

In this section we will present the classification principles derived from solvability conditions of the constraint equations. We will relate them to the several graphs defined before and introduce the reader to the use of the basic graph-theoretic concepts employed in this paper. Since redundancy is defined in terms of observability, the discussion will be focused on observability classification only. Theorems and proofs are given in

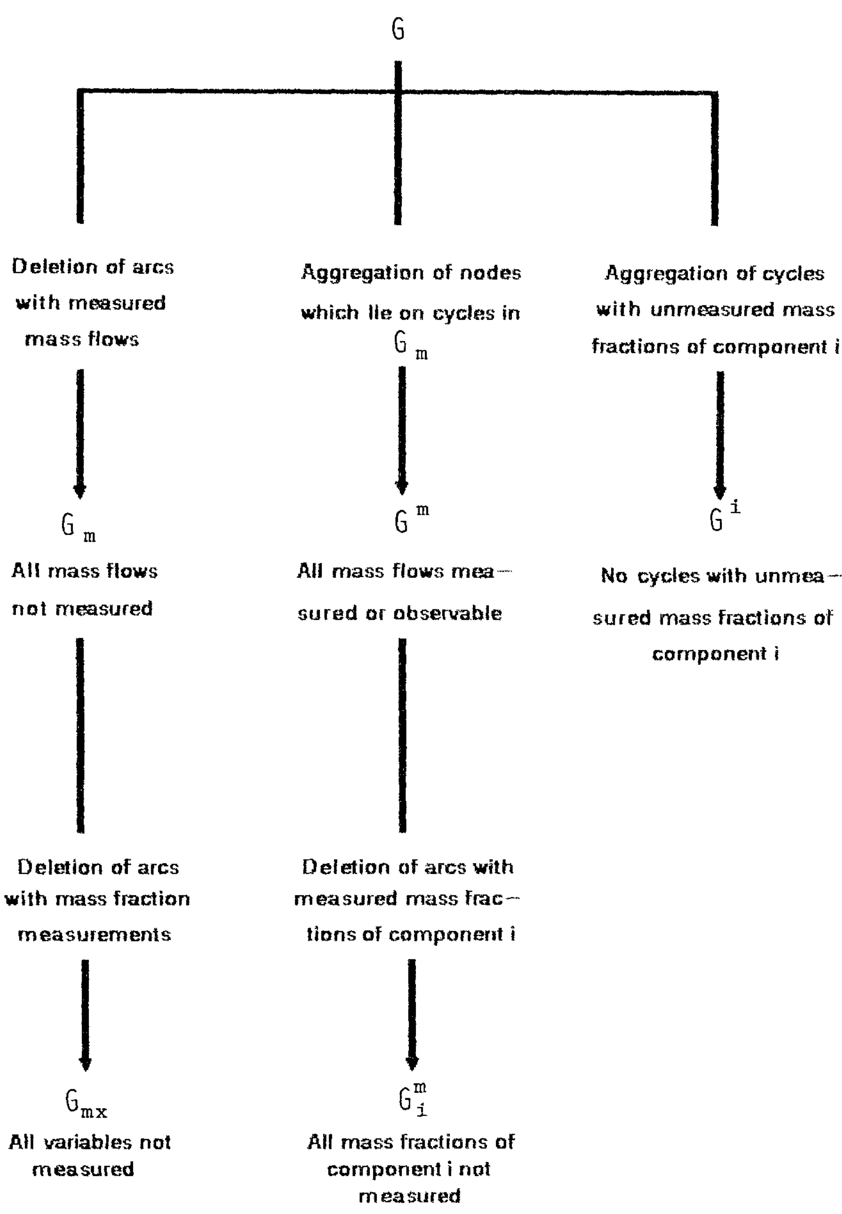

Figure 1. *G* and graphs derived from *G*.

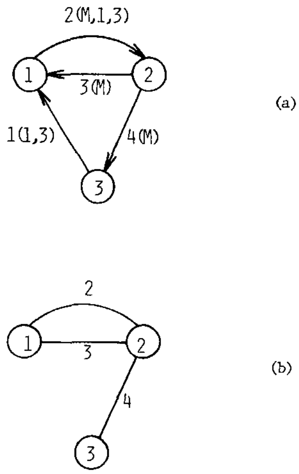

Figure 2. (a) A three-component network.
(b) G_m graph of network.
Quantities in parentheses indicate unmeasured variables: M = unmeasured mass flow; 1, etc. = unmeasured mass fraction of component 1, etc.

Appendix 2 and Supplementary Material No. 2. Let us begin with the relationship between a cutset (defined in Appendix 1) and its associated balance equations. With reference to Figure 6 let S be the set of arcs $\{e_1, e_2, \ldots, e_j, \ldots, e_{J'}\}$ that divides the nodes into two subsets, V_1 and V_2, such that one end node of each arc belongs to V_1 and the other to V_2. Then S is a cutset of graph G. A material balance around the node subset V_1 (or V_2) is the algebraic sum of material flows in the arcs of cutset S. We shall, therefore, use interchangeably the phrases "balance around a node subset" and "balance for a cutset."

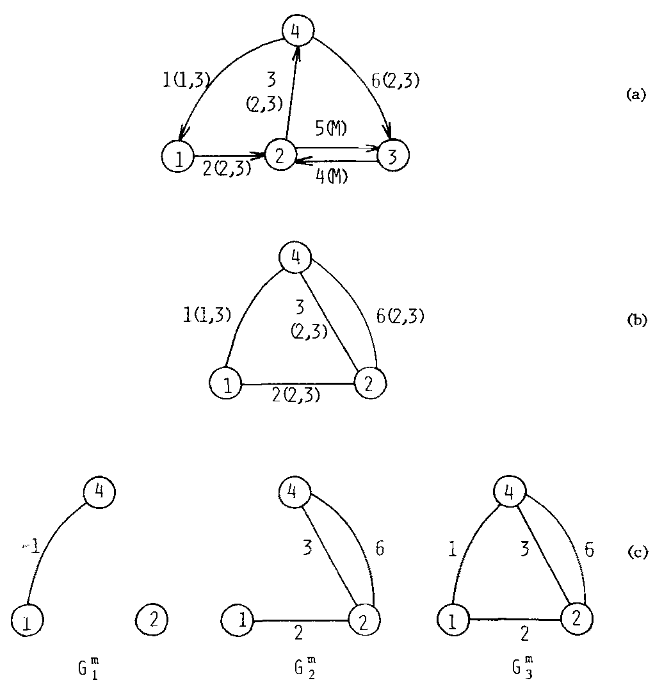

Figure 3. (a) A three-component network.
(b) G^m graph of network.
(c) G_i^m subgraphs of G^m.
Parenthetical designations as in Fig. 2.

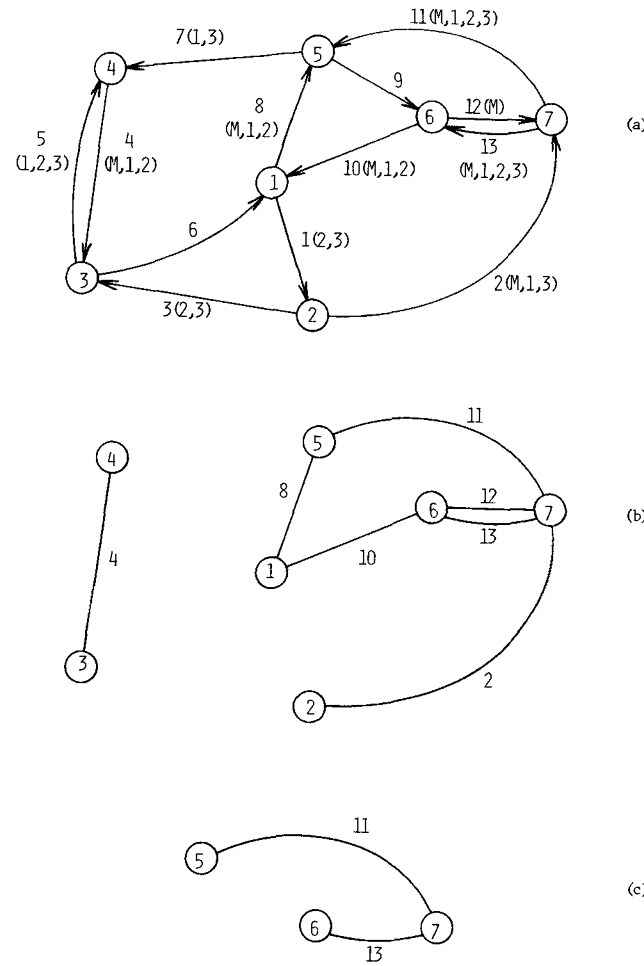

Figure 4. (a) A three-component network.
(b) G_m graph of network.
(c) G_{mx} graph with isolated nodes excluded.
Parenthetical designations as in Fig. 2.

The overall and component balances corresponding to the cutset of G that consists of the arcs in S are:

$$a_1 M_1 + \cdots + a_j M_j + \cdots + a_{J'} M_{J'} = 0 \qquad (4)$$

and

$$a_1 M_1 x_{1i} + \cdots + a_j M_j x_{ji} + \cdots + a_{J'} M_{J'} x_{J'i} = 0 \quad i = 1, \ldots, I \qquad (5)$$

where $a_1, \ldots, a_j, \ldots, a_{J'} = \pm 1$ depending upon the directions of the arcs.

Suppose now that mass flow M_j is not measured, which implies that arc e_j is present in graph G_m. M_j can be calculated through the overall balance if and only if Eq. 4 contains exactly one unknown variable, M_j. In other words, a necessary and sufficient condition for M_j to be observable through the overall balance of some cutset of G is that all arcs in the cutset, except arc e_j, have measured or observable mass flow. Mah et al. (1976) proved that this situation corresponds to the case in which arc e_j does not lie on a cycle in G_m. This condition is sufficient for observability classification of mass flows in multicomponent flow networks and is stated as Theorem 1 (Appendix 2).

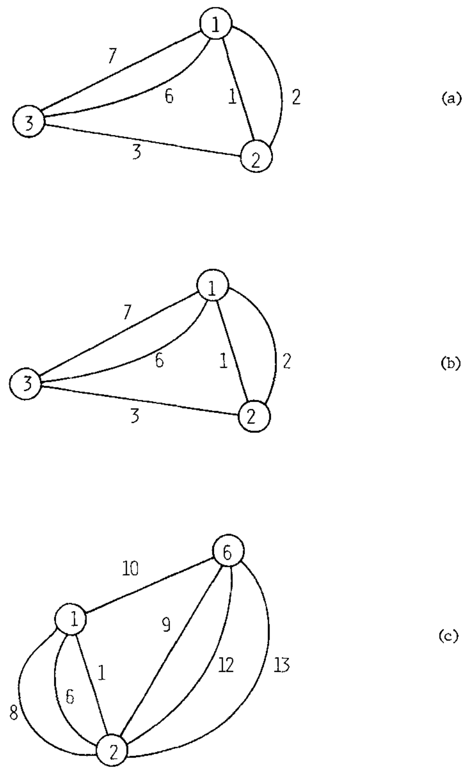

Figure 5. Graphs for the network in Fig. 4a. (a) G^1 graph; (b) G^2 graph; (c) G^3 graph.

The other case to consider is when arc e_j does lie on a cycle with unmeasured mass flows. A property of cutsets is that they have even number of arcs in common with any cycle (Deo, 1974, p. 70). Thus, if arc e_j of set S lies on a cycle of G_m there is at least another arc of S that is part of the same cycle. Then the overall balance, Eq. 4, has more than one unknown mass flow. However, we may still be able to calculate the unknown flow(s) by augmenting Eq. 4 with component balances, Eq. 5. Before analyzing this case we will demonstrate the points made above using the three-component network in Figure 2. Unmeasured flow M_4 can be directly calculated through the overall balance around node 3. Notice that arc 4 does not lie on a cycle in G_m. On the other hand, arcs 2 and 3 do lie on a cycle in G_m. Mass flows

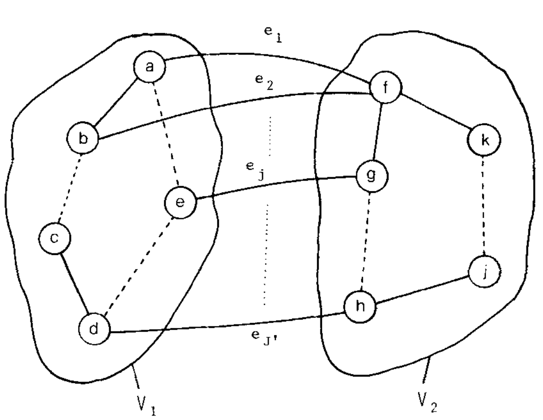

Figure 6. A cutset of G ($e_1, \ldots, e_j, \ldots, e_{J'}$), and partitions V_1, V_2 of node set of G.

M_2 and M_3 may be calculated by using both the overall balance and the material balance for component 2 around node 1.

Let E_m denote the set of arcs that lie on cycles in G_m, and let arc e_j of S together with at least another arc of the same set belong to E_m. Let K_m be the family of cutsets of G each of which contains some arcs of E_m. Clearly, S is a member of K_m. Let $s(s > 1)$ be the number of arcs in S that belong to E_m. If all J' streams of S have measured or observable mass fractions of the same p components ($p \in \{0, 1, \ldots, I\}$), we can define $\underline{x}_k$ as the vector of these mass fractions in arc e_k. The coefficient matrix, $\underline{C}$, of $M_1, M_2, \ldots, M_{J'}$ in the same $p + 1$ equations of the system, Eqs. 4 and 5, is:

$$\underline{C} = \left[a_1 \begin{pmatrix} 1 \\ \underline{x}_1 \end{pmatrix} \vdots \cdots \vdots\, a_j \begin{pmatrix} 1 \\ \underline{x}_j \end{pmatrix} \vdots \cdots \vdots\, a_{J'} \begin{pmatrix} 1 \\ \underline{x}_{J'} \end{pmatrix}\right] \qquad (6)$$

The columns of $\underline{C}$ can be permuted and partitioned so that:

$$\underline{C} = [\underline{C}_s : \underline{C}_q] \qquad (7)$$

where columns of $\underline{C}_s$ correspond to the s arcs that belong to E_m($\underline{m}_s$ being the vector of mass flows) and those of $\underline{C}_q$ correspond to the $q = J' - s$ remaining arcs in S, all of which have measured or observable mass flows ($\underline{m}_q$ being the vector of mass flows). The unmeasured mass flows $\underline{m}_s$ may be determined by solving the following set of equations:

$$\underline{C}_s \underline{m}_s = -\underline{C}_q \underline{m}_q \qquad (8)$$

The necessary and sufficient conditions for a unique solution of Eq. 8 are (Amundson, 1966): rank $[\underline{C}_s]$ = rank $[\underline{C}_s : -\underline{C}_q\underline{m}_q] = s$. In the case when $p = I$, at most I of the $I + 1$ equations in Eq. 8 are linearly independent because $\underline{1}^T\underline{x}_k = 1$, $k = 1, \ldots, J'$. Therefore, for $\underline{m}_s$ to be uniquely determinable, its dimension must not exceed I. Also, for $\underline{m}_s$ to be nonzero Eq. 8 must be nonhomogeneous, i.e., $s < J'$. The sufficient conditions for classifying mass flows of arcs in E_m through some cutset of G are summarized in Theorem 2. Note that when $\underline{C}_s$, $\underline{C}_q$, or $\underline{m}_q$ contain observable unmeasured variables, the above matrix rank test cannot be performed numerically. Then the solvability of Eq. 8 has to be examined symbolically. In other words, we check whether the number of equations is at least equal to s, i.e., $p + 1 > s$. In order to demonstrate how Eq. 8 is set up, let us consider again the three-component flow network in Figure 2a. G_m has one cycle consisting of arcs 2 and 3. The set of cutsets of G containing arcs of $E_m = \{2, 3\}$ is: $K_m = \{(1, 2, 3), (2, 3, 4)\}$. In cutset $K_{m,1} = (1, 2, 3)$ arc 3 has completely measured composition but arcs 1 and 2 have measured mass fractions of component 2 only. Therefore, the partitioned matrix $\underline{C}$ is:

$$[\underline{C}_s | \underline{C}_q] = \begin{bmatrix} -1 & 1 & \vdots & 1 \\ -x_{22} & x_{32} & \vdots & x_{12} \end{bmatrix} \qquad (9)$$

For this system $s = 2$, $I = 3$, and $J' = 3$. If $x_{22} \neq x_{32}$, then M_2 and M_3 are observable.

We now turn our attention to the classification of mass fractions. Suppose x_{ji} is not measured. As mentioned in the previous section, at least another mass fraction x_{jl}, $l \neq i$, is not measured, and so x_{ji} cannot be calculated using the normalization equation for stream e_j. The ith component balance in the set of Eq. 5 is:

$$a_1 M_1 x_{1i} + \cdots + a_j M_j x_{ji} + \cdots + a_{J'} M_{J'} x_{J'i} = 0 \quad (10)$$

Assume that all mass fractions $x_{ki} (e_k \in S, k \neq j)$ are measured and have the same value. Multiplying Eq. 4 by x_{ki} and subtracting it from Eq. 10, we obtain the modified component balance:

$$a_j M_j (x_{ji} - x_{ki}) = 0 \quad (11)$$

Equation 11 implies that $x_{ji} = x_{ki}$, or x_{ji} is observable. Notice that we have not imposed any condition on the observability of mass flows. Since all mass fractions of component i in streams other than e_j are measured, the above classification is possible only if cutset S has no arcs in common with a cycle in which all mass fractions of component i are not measured. Otherwise, there are at least two arcs in S with unmeasured mass fractions of component i.

Let e_j be an arc which has unmeasured x_{ji} and which does not lie on cycles with unmeasured mass fractions of component i, and let E^i be the set of such arcs. The foregoing discussion shows that the modified component balances (11) may be used to classify all unmeasured mass fractions x_{ji}, $e_j \in E^i$. By definition, all members of E^i are contained in G^i. If K_j^i is the set of cutsets of G^i, each of which contains exactly one member of E^i, namely, e_j, then only members of K_j^i have to be checked for observability classification of x_{ji}. The use of G^i greatly simplifies the tasks of identifying x_{ji} and locating the appropriate cutset.

These results are stated in Theorem 3 and its corollary. For an illustration consider the process network in Figure 4a. The associated G^1 graph is given in Figure 5a and $E^1 = \{2, 7\}$. The only cutset in K_2^1 is (1, 2, 3). This cutset can readily be determined if we temporarily aggregate the end nodes of all arcs in E^1 except arc 2 (in this case, end nodes 1 and 3 of arc 7). If $x_{11} = x_{31}$, then x_{21} is observable. Notice that the use of G^1 avoids the necessity of examining all cutsets of G containing arc 2.

In the case when x_{ji} cannot be classified as observable by Theorem 3 or its corollary, we may still be able to calculate it through the ith component balance for some cutset of G. Then with reference to S, Eq. 10 should have exactly one unknown variable, x_{ji}. Let us consider graph G^m. All mass flows in G^m are measured or observable by the definition of G^m and Theorem 1. Consequently, the status of any mass fraction in G^m is the same as of its corresponding component flow. Suppose set S is a cutset of G^m. Then no arc in S belongs to E_m. Rewriting Eq. 10 in terms of component flows we get:

$$a_1 n_{1i} + \cdots + a_j n_{ji} + \cdots + a_{J'} n_{J'i} = 0 \quad (12)$$

Component flow n_{ji} can be calculated from Eq. 12 if and only if all remaining flows n_{ki}, $k \neq j$, are known. Hence, a necessary and sufficient condition for x_{ji} to be observable through some cutset in G^m is that all arcs in the cutset, except arc e_j, have measured or observable flows of component i. Clearly, this case is analogous to the classification of unmeasured mass flows in G through overall balances, which was examined previously. Here, however, we are concerned with G^m instead of G, and G_i^m is the corresponding graph to G_m. By analogy then, if arc e_j does not lie on a cycle in G_i^m, x_{ji} is observable. This result is presented in Theorem 4. An illustration may be found in Figure 3. Mass fraction x_{11} can be estimated in G^m using the balance for component 1 around node 4. Notice that arc 1 lies on no cycle in G_1^m (Figure 3c).

All the classification principles presented so far refer to situations in which unmeasured variables in the arcs of set S are observable. Similarly, there are certain cases in which unmeasured variables in the arcs of S are unobservable. For instance, if $s (s > 1)$ arcs in S belong to E_m and at least one arc in S has unmeasured composition with respect to all components (i.e., $p = 0$), then there are fewer equations in the system of Eq. 8 than unknowns. In this case the unmeasured mass flows in S cannot be calculated through this set of equations. Sufficient conditions for unobservability classification based on this remark are given by Corollaries 1 and 2 of Theorem 2. Also, if an arc in S has unobservable mass flow, no unmeasured mass fraction x_{ji}, $e_j \in S$, can be calculated through cutset S, unless Theorem 3 or its corollary, which classify mass fractions irrespective of the status of mass flows, are applicable—Theorem 5. In the case in which at least two of the same components have no measured mass fractions in a cycle of G, then these mass fractions are unobservable—Theorem 6.

Observability Classification

Algorithm

Given the process network, the measurement placement, and the measurement values the algorithm classifies individually all unmeasured variables as observable or unobservable.

The algorithm may be conceived as consisting of two phases, the pretreatment phase and the main phase. In the pretreatment phase, mass fractions satisfying conditions of Theorem 3 or its corollary, as well as single unmeasured mass fractions that may occur in some streams, are classified as observable. Also, arcs with unobservable variables are identified by applying Corollaries 1 and 2 of Theorem 2, and they subsequently are excluded from the process network by node aggregations. All remaining variables are classified in the second phase of the algorithm, which proceeds iteratively. In order to explain this procedure we introduce the following classification of the unmeasured variables depending on their relation with the cycles of G_m:

- Type I flows: the mass flows of arcs that lie on cycles in G_m
- Type II flows: all remaining unmeasured mass flows (they correspond to single-arc biconnected components of G_m)
- Type I mass fractions: the unmeasured mass fractions of any arc with both end nodes lying on the same cycle in G_m
- Type II mass fractions: all remaining unmeasured mass fractions

Theorem 1 indicates that type II flows can be directly classified as observable. Classification of type I flows makes use of measured as well as observable mass fractions (Theorem 2). On the other hand, classification of type I mass fractions makes use of component mass balances and depends on the classification of type I flows (e.g., Theorem 5). Therefore, the observable mass fractions that are possibly needed for classification of type I flows come from the set of type II mass fractions. This analysis suggests that the classification has to be made according to the following hierarchy:

1. Type II flows are classified as observable
2. As many as possible type II mass fractions are classified as observable
3. Type I flows are classified as far as possible
4. After some type I flows have been classified as observable it may be possible to classify some type I mass fractions

The above procedure is repeatedly applied until no further unmeasured mass flow can be classified as observable.

During the execution of the algorithm, G and its derived graphs are continually changing because of deletion of arcs, aggregation of nodes, and classification of unmeasured variables as observable. In the following presentation it is understood that G and its derived graphs are updated after each and every change, and that whenever an unmeasured variable is classified as observable, it is thereafter treated as measured.

We shall now present the algorithm and then illustrate and explain with more details the various steps by applying it to the three-component flow network in Figure 4a.

1. Obtain graphs G^i and arc sets E^i, $i = 1, \ldots, I$. For each arc j in E^i consider set K_j^i. If in a member of K_j^i there are two arcs or all mass fraction measurements of component i are equal, classify x_{ji} as observable.

2. For each arc j in G with exactly one unmeasured mass fraction x_{ji}, classify x_{ji} as observable.

3. Find components and biconnected components of G_{mx}.

(a) Aggregate in G all nodes of the biconnected components of G_{mx} with more than one arc.

(b) For each arc j in G_m with mass fraction measurements and with end nodes v and w lying in the same component of G_{mx}, aggregate nodes v and w in G and also all nodes in the cycle of G formed by arc j and arcs in G_{mx}.

(c) Classify all unmeasured variables of the arcs excluded from G by node aggregations in steps (a) and/or (b) as unobservable.

4. Find biconnected components of G_m. For each biconnected component with exactly one arc j, classify M_j as observable and delete arc j from G_m.

5. Obtain graphs G_i^m ($i = 1, \ldots, I$) and form set Z of their biconnected components.

(a) For each biconnected component $B_{i,k}^m$ with exactly one arc j, classify x_{ji} as observable and delete graph $B_{i,k}^m$ from set Z.

(b) For each arc j that appears only once in set Z, say in $B_{t,r}^m$, classify x_{jt} as observable and delete arc j from graph $B_{t,r}^m$.

(c) Update set Z by finding the biconnected components of the graphs $B_{t,r}^m$ which changed in step (b) and repeat from (a) until no further mass fraction is classified as observable.

6. If G_m is a null graph, classify all unclassified unmeasured mass fractions as unobservable and stop.

7. Consider the set of cutsets K_m and one of its members $K_{m,r}$.

(a) If all arcs in $K_{m,r}$ have unmeasured unclassified mass flows, or if the number of arcs in $K_{m,r}$ is greater than I, consider another cutset $K_{m,t} (t \neq r)$ and repeat the step.

(b) If in a cutset $K_{m,r}$ there are s arcs $j1, j2, \ldots, js$ with unmeasured unclassified mass flows and matrices $\underline{C}_s$ and $[\underline{C}_s : -\underline{C}_q \underline{m}_q]$ are of rank s, classify $M_{j1}, M_{j2}, \ldots, M_{js}$ as observable. Delete arcs $j1, j2, \ldots, js$ from G_m and go to step 4.

(c) If after a complete pass through the set of cutsets K_m no mass flow has been classified as observable, classify all unclassified variables (mass flows and mass fractions) as unobservable and stop.

Example

1. In the first step of the algorithm we apply Theorem 3 and its corollary. The mass fractions and cutsets to be considered are identified from graphs G^i and arc sets E^i, $i = 1, \ldots, I$. For the example the G^i graphs are shown in Figure 5. The sets E^i are $E^1 = \{2, 7\}$, $E^2 = \{1, 3\}$, $E^3 = \{1, 13\}$. It is easy to verify that no mass fraction can be classified as observable by Theorem 3 because no cutset in a set $K_j^i (j \in E^i, i = 1, 2, 3)$ has exactly two arcs. However, the application of its corollary may result in classifying some mass fractions as observable. For instance, in G^1 if $x_{31} = x_{61}$, then x_{71} is observable. We assume, however, that no two mass fraction measurements have the same value, and so no mass fraction is classified as observable in this step.

2. In step 2 a single unclassified unmeasured mass fraction that may occur in some stream in the original network or after step 1, is classified as observable since it can be estimated through the corresponding normalization equation. In the example, no stream has exactly one unmeasured mass fraction. After this step, all partially measured stream compositions have at least two unclassified unmeasured mass fractions. Since Theorem 3 and its corollary have been applied in step 1, we have proceeded as far as possible in classifying unmeasured mass fractions as observable. We shall now apply Theorem 5 to classify other mass fractions as unobservable.

3. In step 3(a) cycles of G_{mx} are identified through its biconnected components with more than one arc. All arcs lying in such cycles have unobservable mass flows by Corollary 1 of Theorem 2. In the example, G_{mx} in Figure 4c has no such biconnected component, and the network remains unaltered.

If the end nodes v and w of an arc j, with mass fraction measurements, lie on the same component of G_{mx}, then there is at least one path between v and w that consists of unmeasured arcs. Furthermore, if M_j is unmeasured (i.e., arc j exists in G_m), then a cycle is formed by arc j and arcs in G_{mx} where exactly one arc has mass fraction measurements (arc j). By Corollary 2 of Theorem 2 all mass flows in such a cycle are unobservable. In the example, arc 12 exists in G_m (see Figure 4b) but not in G_{mx}, and also its end nodes, 6 and 7, lie on the same component of G_{mx}. Therefore, arcs 12 and 13 form a cycle satisfying the conditions of Corollary 2 of Theorem 2, and we aggregate nodes 6 and 7 into node 6.

In step 3(c) we classify all unmeasured mass flows of the arcs excluded from G in both previous steps as unobservable. Also, applying Theorem 5 we classify all unmeasured mass fractions in the same arcs as unobservable. In the example, M_{12}, M_{13}, $x_{13,1}$, $x_{13,2}$, and $x_{13,3}$ are classified as unobservable. The new G is given in Figure 7.

4. In step 4 we apply Theorem 1. In the example, arcs 2 and 4 form single-arc biconnected components of G_m, Figure 8. We classify M_2 and M_4 as observable and delete these arcs from G_m. After this step all biconnected components of G_m consist of more than one arc. For the example, the new G_m graph is the same as $B_{m,1}$ in Figure 8.

5. In step 5 type II mass fractions are classified as observable by applying Theorem 4. In the original set Z of biconnected components of $G_i^m (i = 1, \ldots, I)$ all arcs appear at least twice and as many times as the number of their unclassified unmeasured mass fractions. For the example, this set is given in Figure 9.

In accordance with condition (a) of Theorem 4 an unmeasured mass fraction x_{ji} is observable if arc j forms a single-arc biconnected component of G_i^m. In the example, graphs $B_{1,2}^m$, $B_{1,3}^m$, $B_{2,2}^m$ and $B_{2,3}^m$ are single-arc graphs in Z and, consequently, x_{21},

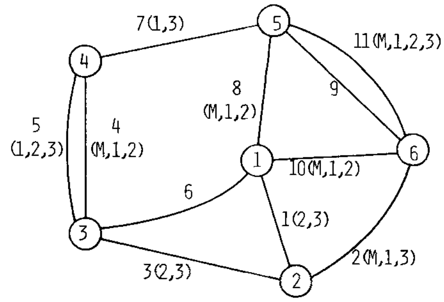

Figure 7. Revised graph G after step 3 of observability algorithm.
Parenthetical designations as in Fig. 2.

x_{71}, x_{12}, and x_{32} are classified as observable. Also, all these graphs are deleted from set Z.

After step 5(a) there may be arcs appearing only once in set Z, and therefore all their mass fractions except one are measured or observable. Now we can classify the last unclassified mass fraction of such arcs as observable, since it can be estimated through the normalization equation. In the example, x_{13}, x_{23}, x_{33} and x_{43} are classified as observable because, after step 5(a) arcs 1, 2, 3, and 7 appear only once in Z. Furthermore, these arcs are deleted from graph $B^m_{3,1}$.

Because of possible deletions of arcs in step 5(b), cycles in the original set Z may no longer exist; therefore, we can again apply conditions (a) and (b) of Theorem 4. Repeating steps 5(a) and 5(b) in the example, x_{53} is classified as observable. The revised G is shown in Figure 10. A comparison of Figures 4a and 10 shows the reduction in the number of unclassified variables.

6. If G_m is a null graph, then all unmeasured mass flows have been classified and G^m is identical with G. Since Theorem 4, which provides necessary and sufficient conditions for observability classification of mass fractions in G^m, has been applied in step 5, all unclassified mass fractions are unobservable and the algorithm stops. In the example, G_m is not a null graph and we proceed to the next step.

7. In step 7 we consider set K_m, which consists of cutsets of G containing arcs of G_m. In steps (a) and (b), the conditions of Theorem 2 are checked for cutsets of K_m until the mass flows of some arcs in G_m are classified as observable. Then these arcs are deleted from G_m and we return to step 4, since graphs G_m, G^m, and G have been modified. Examination of all cutsets of K_m constitutes a sufficient and necessary condition for classifying all unclassified unmeasured mass flows as unobservable. In this case, type I mass fractions can be classified as unobservable according to Theorem 5. Also, all type II mass fractions can be classified as unobservable because they were not classified as observable in G^m (step 5), which includes all balances of G with measured or observable mass flows, while all remaining balances in G but not in G^m contain unobservable mass flows. In the example, G_m consists of one biconnected component ($B_{m,1}$ in Figure 8) and $E_m = \{8, 10, 11\}$. If we consider cutset $K_{m,1} = (1, 6, 8, 10)$ of G in Figure 10, then, assuming $x_{83} \neq x_{10,3}$, we classify M_8 and M_{10} as observable, since $s = 2$, $J' = 4$, $I = 3$, and all mass fractions of component 3 are measured ($p = 1$). We delete arcs 8 and 10 from G_m and return to step 4.

The new G_m graph consists of just one arc, arc 11, and M_{11} is classified as observable in step 4. At this point we have completed the classification of mass flows; deleting arc 11 from G_m, this graph becomes null. Considering the new set Z in step 5, $x_{11,3}$ is classified as observable. Finally in step 6, the algorithm stops because G_m is a null graph. All remaining mass fractions x_{41}, x_{42}, x_{51}, x_{52}, x_{81}, x_{82}, $x_{10,1}$, $x_{10,2}$, $x_{11,1}$, and $x_{11,2}$ are unobservable.

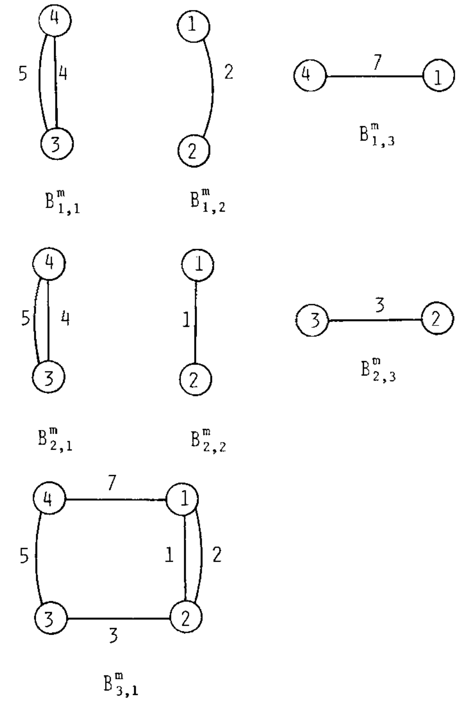

Figure 9. Original set Z for graph G in Fig. 7.

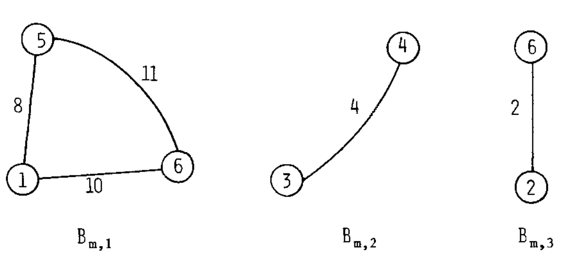

Figure 8. Biconnected components of G_m for graph G in Fig. 7.

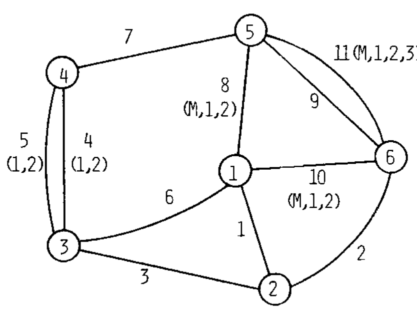

Figure 10. Revised graph G after step 5 of observability algorithm.
Parenthetical designations as in Fig. 2.

Redundancy Classification

Algorithm

Since a measurement is redundant if its deletion causes no loss of observability, the redundancy classification of a measured variable can be determined by deleting the measurement and applying the observability classification algorithm. Depending on whether the variable is observable or unobservable, the corresponding measurement is classified as redundant or nonredundant. However, such a procedure is not very efficient because we have to delete each measurement in turn and apply the observability algorithm as many times as the number of measurements. In particular, the large number of variables and measurements in multicomponent flow networks dictates that this "brute force" procedure should be avoided except as a last resort. For this reason, classification rules have been developed, according to which measurements are directly classified without applying the observability algorithm. These rules have been derived by using the definition of redundancy and the observability classification theorems. They are stated in Theorems 7 to 12, which are presented together with illustrations in Appendix 2.

In addition to the redundancy theorems, the algorithm to be presented directly classifies measurements in streams with completely measured composition as well as in streams excluded from the process network by applying steps 3(a) and 3(b) of the observability algorithm. In the former type of streams, all mass fraction measurements are redundant through the normalization equation. On the other hand, the mass flow measurements in the streams obliterated in step 3 of the observability algorithm are nonredundant by Theorem 8, while the mass fraction measurements of these streams with completely measured compositions are redundant, as mentioned before. The only measurements remaining for classification are the mass fraction measurements in the streams with at least one unmeasured mass fraction. If we delete any of these measurements and apply step 3 of the observability algorithm, the corresponding variable will be classified as unobservable. Hence, all these mass fraction measurements are nonredundant. Note that no mass fraction measurement can be classified as redundant by Theorem 9 because all cutsets of G containing arcs excluded from G in steps 3(a) and 3(b) of the observability algorithm have at least one arc corresponding to a stream with mass fraction measurements.

We shall now present the algorithm, the various steps of which will be illustrated and explained with respect to the three-component flow network in Figure 11a.

1. Find components of G_m. For each arc j with end nodes in different components of G_m classify measurement M_j as redundant.

2. Find components of G_{mx}. For each arc j with measured mass flow and end nodes in the same component of G_{mx} classify measurement M_j as nonredundant.

3. For each stream j with completely measured composition classify all measurements x_{ji} $(i = 1, \ldots, I)$ as redundant.

4. Apply steps 3(a) and 3(b) of the observability classification algorithm. Classify all unclassified mass fraction measurements in the streams excluded from G as nonredundant.

5. Obtain graphs G^i, $i = 1, \ldots, I$. In each graph G^i aggregate the end nodes of all arcs with unmeasured mass fractions of ith component. Classify measurement x_{ji} as redundant if arc j exists

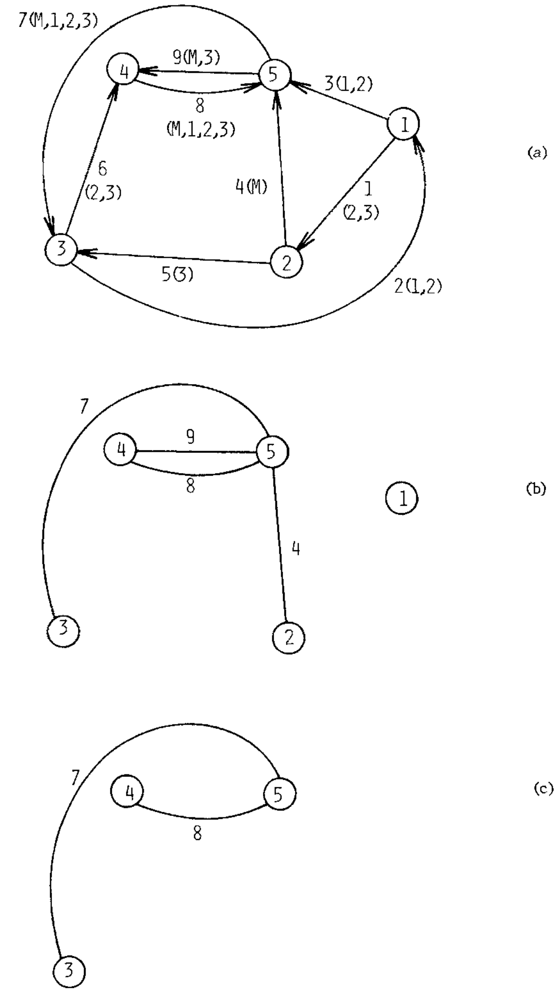

Figure 11. (a) A three-component network.
(b) G_m graph of network.
(c) G_{mx} graph with isolated nodes excluded.
Parenthetical designations as in Fig. 2.

in a cutset of G^i in which all mass fraction measurements of ith component have the same value.

6. Obtain graphs G_i^m $(i = 1, \ldots, I)$ and find their components and biconnected components. Classify measurement x_{ji} as redundant if:

(a) end nodes of arc j lie on different components of G_i^m, or

(b) arc j is a single-arc biconnected component in all other graphs $G_t^m (t \neq i, t = 1, \ldots, I)$.

7. Classify measurement x_{ji} as nonredundant if:

(a) there is an alternative path P between end nodes of arc j and on this path mass fractions of component i are not measured, and if

(b) there is at least one component k, $k \neq i$, with unmeasured mass fractions in the cycle formed by arc j and path P.

8. Delete temporarily each unclassified measurement in turn and apply the observability classification algorithm. If the variable is classified as observable (respectively, unobservable) classify the measurement as redundant (respectively, nonredundant).

It is important to note that in the case of multiple measurements (i.e., variables measured more than once), these measurements should be directly classified as redundant without applying the algorithm.

Example

1. The redundancy classification of mass flow measurements in step 1 is based on Theorem 7. Measurements M_1, M_2, and M_3 are redundant because arcs 1, 2, and 3 are incident to node 1, which constitutes a component of G_m, Figure 11b.

2. The classification in step 2 is a direct application of Theorem 8. In the example, this step results in classifying measurement M_6 as nonredundant since end nodes of arc 6 lie in the same component of G_{mx}, Figure 11c.

3. Mass fraction measurements of streams completely measured with respect to composition are classified as redundant in step 3. In the example, stream 4 is such a stream, and x_{41}, x_{42}, x_{43} are classified as redundant.

4. At this point we apply steps 3(a) and 3(b) of the observability algorithm. The mass flow measurements in the arcs excluded from G have been classified in step 2 and the mass fraction measurements of streams with completely measured composition in step 3. All remaining measurements in the streams excluded from G are classified as nonredundant according to what was mentioned before presenting the algorithm. In the example, G is reduced only in step 3(b) of the observability algorithm. Arcs 8 and 9 form a cycle in which one arc has mass fraction measurements (arc 9) and nodes 4 and 5 are aggregated into node 4. Measurements x_{91} and x_{92} are classified as nonredundant.

5. In step 5 we apply Theorem 9. No cycle of arcs with unmeasured mass fractions of ith component exists in G^i. Furthermore, by aggregating in G^i the end nodes of all remaining arcs with unmeasured mass fractions of the ith components, the cutsets of the graph so derived (updated G^i) are those cutsets of G in which all mass fractions of the ith component are measured. If an arc j exists in a cutset of the updated G^i in which all mass fraction measurements of the ith component have the same value, then measurement x_{ji} is redundant according to Theorem 9. For the example, the updated G^i graphs are given in Figure 12. Notice that the updated G^2 and G^3 are single-node graphs. In G^1, if $x_{11} = x_{41} = x_{51}$, then these measurements are redundant. We assume, however, that no two measurements have the same value, and so no classification is performed at this step.

6. Mass fraction measurements satisfying the conditions of Theorems 10 and 11 are classified as redundant in steps 6(a) and 6(b), respectively. For the example, G^m and its subgraphs $G_i^m (i = 1, 2, 3)$ are given in Figure 13. Note that G^m is the same as G after step 4. In step 6(a) measurements x_{11} and x_{51} are classified as redundant because the end nodes of arcs 1 and 5 lie on different components of G_1^m(x_{41} has been classified in step 3). Measurement x_{52} is classified as redundant in step 6(b) because arc 5 forms a single-arc biconnected component in G_3^m (i.e., it does not lie on a cycle in this graph).

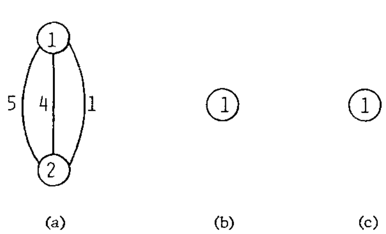

Figure 12. Updated G^i graphs for the network in Fig. 11a.
(a) G^1 graph; (b) G^2 graph; (c) G^3 graph

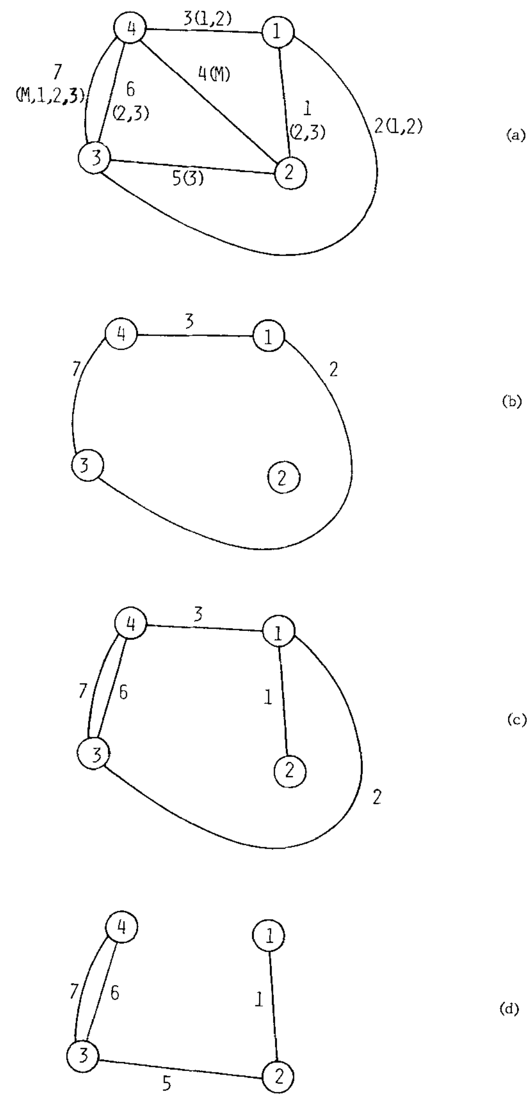

Figure 13. Graphs for the network in Fig. 11a.
(a) G^m graph; (b) G_1^m subgraph; (c) G_2^m subgraph; (d) G_3^m subgraph.

7. In step 7 mass fraction measurements are classified as nonredundant by applying Theorem 12. For an arc j with an unclassified measurement x_{ji}, there is at least one component k, $k \neq i$, with unmeasured x_{jk}. Otherwise, x_{ji} would have been classified as redundant in step 3. In addition, if there is an alternative path P between the end nodes of arc j and on this path all mass fractions of components i and k are not measured, then a cycle c is formed by arc j and path P satisfying the conditions of Theorem 12. For instance, consider measurement x_{61}. In graph G of Figure 13a the end nodes of arc 6 are connected by a path (arc 7) along which mass fractions of component 1 are not measured. Also, in the cycle (6, 7) component 2 (and 3) has unmeasured mass fractions. Hence, measurement x_{61} is nonredundant.

8. At this point, all the graph-theoretic criteria for redundancy classification have been applied. The definition of redundancy and the observability algorithm have to be employed for classification of the remaining unclassified measurements M_5, x_{23}, and x_{33}. Deleting each measurement in turn and applying the observability algorithm, all measurements are classified as redundant.

Closing Remarks

One advantage of the proposed algorithms is that we are able to classify every variable and measurement, which was not always possible with the previous algorithms (Stanley and Mah, 1981b). Also, the approach presented in this paper provides a framework for generalizing the treatment to include chemical reactions and energy flows. This extension is expected to be nontrivial. However, based on our past experience we are optimistic of the outcome. The availability of observability and redundancy algorithms will provide a basis and means for designing performance monitoring systems.

Acknowledgment

This work was supported by National Science Foundation Grant No. CBT-8519182 and by the award of an IBM Graduate Predoctoral Fellowship to A. Kretsovalis.

Notation

$\underline{A}$ = reduced incidence matrix of graph F
a_{nj} = (nj)th entry of $\underline{A}$
B_m = set of biconnected components of G_m; $B_m = \{B_{m,1}, B_{m,2}, \ldots\}$
B_i^m = set of biconnected components of G_i^m; $B_i^m = \{B_{i,1}^m, B_{i,2}^m, \ldots\}$
c = cycle in graph G
$\underline{C}$ = mass flow coefficient matrix, Eq. 6
$\underline{C}_s$, $\underline{C}_q$ = submatrices of $\underline{C}$, Eq. 7
E_m = set of arcs that lie on cycles in G_m
E^i = set of arcs in G^i with unmeasured mass fraction of component i
F = a process graph
G = underlying graph of F with directions of arcs erased
G_m = subgraph of G with deletion of all arcs with mass flow measurements
G_{mx} = subgraph of G_m with deletion of all arcs with any mass fraction measurements
G^i = graph obtained from G by aggregating the end nodes of all arcs lying on cycles with unmeasured mass fractions of component i
G^m = graph obtained from G by aggregating the end nodes of all arcs belonging to set E_m
G_i^m = subgraph of G^m with deletion of all arcs with mass fraction measurements of component i
i = general subscript for components
I = number of components
j = general subscript for arcs (streams)
J = number of process streams (arcs)
J' = number of arcs in a member of set K_m
K_m = set of cutsets of G that contain arcs of set E_m; $K_m = \{K_{m,1}, K_{m,2}, \ldots\}$
K_j^i = set of cutsets of G^i each member of which contains only arc j from set E^i
$\underline{m}_s$ = vector of mass flows in arcs belonging to set E_m in a member of K_m
$\underline{m}_q$ = vector of measurd or observable mass flows in a member of K_m
M_j = mass flow in arc j
n_{ji} = mass flow of component i in arc j
N = number of physical nodes in graph F
p = number of components that have measured or observable mass fractions in a member of K_m
P = path in graph G
s = number of arcs belonging to set E_m in a member of K_m
v or w = a node
x_{ji} = mass fraction of component i in stream j
$\underline{x}_j$ = vector of p measured or observable mass fractions in stream j
Z = set of biconnected components of graphs G_i^m, $i = 1, \ldots, I$

Appendix 1: Graph-Theoretic Terminology

An undirected (resp., directed) graph $G = (V, E)$ consists of a set of objects $V = \{v_1, v_2, \ldots\}$ called nodes or vertices, and another set $E = \{e_1, e_2, \ldots\}$ whose elements are called arcs or edges, such that each arc e_k is identified with an unordered (resp., ordered) pair (v_i, v_j) of nodes. The nodes v_i, v_j associated with arc e_k are called the end nodes of e_k. A graph that consists only of vertices is referred to as a null graph. However, a graph can never contain only edges without their end vertices. This possibility is excluded by the formal definition of an edge.

A path is defined as a sequence of nodes and arcs, beginning and ending with nodes, in which no node appears more than once. A cycle is a closed path where only the beginning (or ending) node appears twice. A graph is connected if there is a path between every pair of nodes. A graph g is said to be a subgraph of a graph G if all the nodes and arcs of g are in G. A component of a graph G is a maximal connected subgraph of G. A cutset of a connected graph is a minimal set of arcs whose removal leaves the graph disconnected. In the connected graph of Figure A1-1a the sequence of nodes and arcs $a1b2c$ forms a path between nodes a and c, while the sequence $b2c3d4b$ forms a cycle. Also, the set of arcs $\{2, 3\}$ is a cutset, whereas the set $\{1, 2, 4\}$ is not, because it is not minimal (its proper subsets $\{1\}$, $\{2, 4\}$ are cutsets of G).

A node v is a cutnode of a connected graph G, if its removal together with its incident arcs disconnects G. A graph is separable if it contains a cutnode. If we split a cutnode into two nodes, two disjoint subgraphs are produced. If we repeat this operation until all subgraphs are nonseparable, then the resulting subgraphs are called blocks or biconnected components. The graph in Figure A1-1a contains one cutnode, node b, and has two biconnected components, shown in Figure A1-1b. Two arcs belong to the same biconnected component if and only if they belong to a common cycle. A single arc that does not lie on any cycle forms its own biconnected component (e.g., arc 1 in the example). Any two biconnected components are either disjoint

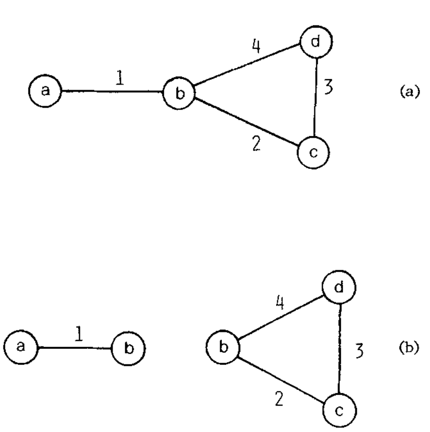

Figure A1-1. (a) A graph.
(b) Biconnected components of the graph.

or have exactly one cutnode in common. Each cutnode lies in at least two biconnected components and all other nodes can belong to only one biconnected component.

Appendix 2: Classification Theorems

Observability classification

Theorem 1. If an arc j with unmeasured mass flow constitutes a single-arc biconnected component of G_m, then M_j is observable.

Proof. It follows directly from the discussion in the Classification Principles section and the property of biconnected components: an arc forms its own biconnected component if and only if it does not lie on any cycle.

Theorem 2. For a cutset $K_{m,r}$ with J' arcs, let $\underline{m}_s$ be the vector of unmeasured mass flows of arcs in E_m and $\underline{m}_q$ the vector of observable mass flows in the remaining arcs of $K_{m,r}$, and let $\underline{C}_s$, $\underline{C}_q$ be their coefficient matrices in the balance equations. Then the mass flows $\underline{m}_s$ are observable, if:

(a) $s \leq I$

(b) $s < J'$, and

(c) rank $[\underline{C}_s]$ = rank $[\underline{C}_s \vdots -\underline{C}_q \underline{m}_q] = s$.

Corollary 1. If arc j lies on a cycle of G_{mx}, then M_j is unobservable.

Proof. Since G_{mx} is a subgraph of G_m, arc j lies on a cycle in G_m. Consequently, any cutset of G containing arc j has at least two unmeasured mass flows ($s \geq 2$), and $p = 0$ because at least two arcs of the cutset are unmeasured. Therefore, condition (c) of Theorem 2 can never be satisfied, and M_j is unobservable.

Corollary 2. If arc j lies on a cycle of G_m with exactly one arc with mass fraction measurements, then M_j is unobservable.

Proof. Similar to that of Corollary 1 with only one difference: $p = 0$ for all cutsets of G containing arc j because they contain at least one unmeasured arc.

Illustrations of the above corollaries may be found in Figure A2-1. Arcs 2 and 3 form a cycle in G_{mx} (Figure A2-1c). Hence, M_2 and M_3 are unobservable (Corollary 1). Also, arcs 4 and 5 form a cycle in G_m (Figure A2-1b) with exactly one arc with

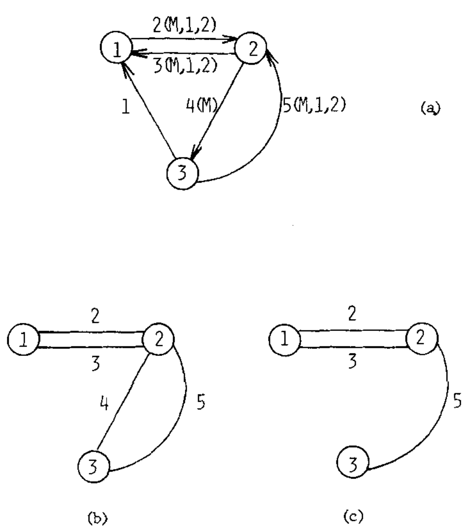

Figure A2-1. (a) A two-component network.
(b) G_m graph of network.
(c) G_{mx} graph.

mass fraction measurements (arc 4). Therefore, M_4 and M_5 are unobservable (Corollary 2).

Theorem 3. If arcs $j1$ and $j2$ form a cutset of $G^i (i = 1, \ldots, I)$ and $x_{j1,i}$ is measured, then $x_{j2,i}$ is observable.

Corollary. Suppose arcs $j1, j2, \ldots, jr$ form a cutset of $G^i (i = 1, \ldots, I)$ and $x_{j2,i}, x_{j3,i}, \ldots, x_{jr,i}$ are measured. If $x_{j2,i} = x_{j3,i} = \ldots = x_{jr,i}$, then $x_{j1,i}$ is observable.

The proofs of both previous results follow directly from the discussion in the "Classification Principles section."

Theorem 4. Unmeasured mass fraction x_{ji} of an arc j in G^m is observable if and only if either of the following conditions is satisfied:

(a) arc j does not lie on a cycle in G_i^m, or

(b) all other mass fractions of arc j are measured or observable.

The proof is given in Supplementary Material No. 2.

For a complete illustration of the above theorem consider the three-component network in Figure 3. Mass fractions x_{11} and x_{22} are observable because arcs 1 and 2 do not lie on a cycle in G_1^m and G_2^m, respectively [condition (a)]. Then x_{13} and x_{23} are observable through the normalization equations of streams 1 and 2 because all remaining mass fractions of these streams are measured (x_{12}, x_{21}) or observable (x_{11}, x_{22}) [condition (b)].

Note that a mass fraction classified as observable in G^m is also observable in G since the constraint equations in G^m constitute a subset of the constraint equations in G. This is not valid, however, for unobservability classification because a mass fraction classified as unobservable by Theorem 4 in G^m might be observable in G through constraint equations corresponding to G but not to G^m.

Theorem 5. If the end nodes of an arc j with at least two missing mass fraction measurements lie on a cycle of arcs with unobservable mass flows and Theorem 3 or its corollary are not applicable, then all unmeasured x_{ji}, $i = 1, \ldots, I$, are unobservable.

Proof. Unmeasured mass fractions x_{ji}, $i = 1, \ldots, I$, cannot be estimated through the normalization equation since more than one mass fraction measurement is missing, or through any modified component balance since Theorem 3 and its corollary cannot be applied. Hence, component mass balances and mass flows are needed. However, all balances including arc j contain unobservable mass flows. Therefore, no unmeasured x_{ji} can be estimated. So, all of them are unobservable.

Note that arc j may have measured or unobservable mass flow. In the latter case it belongs to the cycle.

For demonstrating an application of the above theorem, consider the example for Corollary 1 of Theorem 2 (Figure A2-1). M_2, M_3 are unobservable and arcs 2 and 3 form a cycle in G. Therefore, x_{21}, x_{22}, x_{31}, and x_{32} are unobservable also.

Theorem 6. Let c be a cycle in G and let each of the arcs in c have unmeasured mass fractions for components $i1, i2, \ldots, ik$, $k > 1$. Then all these mass fractons are unobservable.

The proof is given in Supplementary Material No. 2.

For an illustration of the above theorem consider the cycle formed by arcs 4 and 5 in Figure 4a. Mass fractions of components 1 and 2 are not measured, and therefore x_{41}, x_{42}, x_{51}, and x_{52} are unobservable.

It is worth noticing that if cycle c exists in both G and G^m (i.e., all mass flows of arcs in c are measured or observable), Theorems 4 and 6 are equivalent for unobservability classification. Even though Theorem 6 is not explicitly employed in the observability algorithm, it is used for deriving a graph-theoretic rule for redundancy classification.

Redundancy classification

The proofs of all following theorems may be found in Supplementary Material No. 2.

Theorem 7. If the end nodes of an arc j with measured mass flow lie on different components of G_m, measurement M_j is redundant (Theorem 1).

Note that Theorem 7 has been developed on the basis of Theorem 1 and for this reason the latter is mentioned. Similar references will be made in all subsequent theorems.

Theorem 8. If the end nodes of an arc j with measured mass flow lie on the same component of G_{mx}, measurement M_j is nonredundant (Corollaries 1 and 2 of Theorem 2).

Illustrations of the above theorems may be found in Figure 4. The end nodes of arcs 3, 6, and 7 lie on different components of G_m (shown in Figure 4b) and measurements M_3, M_6, and M_7 are redundant (Theorem 7). On the other hand, the end nodes of arc 9 (nodes 5 and 6) lie on the same component of G_{mx} (shown in Figure 4c), and measurement M_9 is nonredundant (Theorem 8).

Theorem 9. Let mass fractions of the ith component be measured in a cutset of G^i, $i \in \{1, \ldots, I\}$, which consists of arcs $j1$, $j2, \ldots, jr$. If $x_{j1,i} = x_{j2,i} = \ldots = x_{jr,i}$, then all these measurements are redundant (Theorem 3 and its corollary).

Theorem 10. If the end nodes of an arc j lie on different components of G_i^m, $i \in \{1, \ldots, I\}$, then measurement x_{ji} is redundant [condition (a) of Theorem 4].

Theorem 11. Let mass fraction x_{ji} be measured and end nodes of arc j belong to the same component of G_i^m, $i \in \{1, \ldots, I\}$. If arc j does not lie on a cycle in any other graph $G_t^m (t \neq i;\ t = 1, \ldots, I)$, then x_{ji} is redundant [condition (b) of Theorem 4].

One difference between Theorem 9 and the two last theorems is that Theorem 9 makes use of measurement values whereas Theorems 10 and 11 are solely graph-theoretic rules. For instance in the network of Figure 3a, measurements x_{21}, x_{31}, and x_{61} may be classified as redundant according to Theorem 9 only if their values are the same (arcs 2, 3, and 6 form a cutset in G and in G^1, which is identical with G). On the other hand, the same measurements are classified as redundant by applying Theorem 10 to G_1^m (Figure 3c), irrespective of their values. Another difference is that Theorem 9 does not stipulate any restriction on the nature of mass flows whereas Theorems 10 and 11 refer to graph G^m, all mass flows in which are measured or observable (by Theorem 1). In the network of the example used before (see Figure 3a), x_{41} and x_{51} may be classified as redundant through a cutset of G^1 (arcs 4, 5, and 6) by applying Theorem 9. However, arcs 4 and 5 do not exist in G^m (they lie on a cycle in G_m). Hence, classification of the above measurements is not possible by Theorems 10 and 11.

Theorem 12. Let x_{ji} be measured and arc j lie on a cycle c of G where mass fractions of component $k (k \neq i)$ are not measured. If all arcs, except arc j, in cycle c have unmeasured mass fractions of component i then measurement x_{ji} is nonredundant (Theorem 6).

An illustration of the above theorem may be found in Figure 4a. Arc 4 lies on cycle $c = (4, 5)$ where mass fractions of component 1 are not measured. In addition, arc 5 has an unmeasured mass fraction of component 3. Therefore, measurement x_{43} is nonredundant.

Literature cited

Aho, A. V., J. E. Hopcroft, and J. D. Ullman, *The Design and Analysis of Computer Algorithms*, Addison-Wesley, Reading, MA (1974).

Amundson, N. R., *Mathematical Models in Chemical Engineering*, Prentice-Hall, Englewood Cliffs, NJ (1966).

Crowe, C. M., Y. A. Garcia Campos, and A. Hrymak, "Reconciliation of Process Flow Rates by Matrix Projection. I: The Linear Case," *AIChE J.*, **29**, 881 (1983).

Deo, N., *Graph Theory, with Applications to Engineering and Computer Science*, Prentice-Hall, Englewood Cliffs, NJ (1974).

Mah, R. S. H., G. M. Stanley, and D. M. Downing, "Reconciliation and Rectification of Process Flow and Inventory Data," *Ind. Eng. Chem. Process Des. Dev.*, **15**, 175 (1976).

Romagnoli, J. A., and G. Stephanopoulos, "On the Rectification of Measurement Errors for Complex Chemical Plants," *Chem. Eng. Sci.*, **35**, 1067 (1980).

Stadtherr, M. A., W. A. Gifford, and L. E. Scriven, "Efficient Solution of Sparse Sets of Design Equations," *Chem. Eng. Sci.*, **29**, 1025 (1974).

Stanley, G. M., and R. S. H. Mah, "Observability and Redundancy in Process Data Estimation," *Chem. Eng. Sci.*, **36**, 259 (1981a).

———, "Observability and Redundancy Classification in Process Networks," *Chem. Eng. Sci.*, **36**, 1941 (1981b).

Vaclavek, V., "Studies on System Engineering. III: Optimal Choice of the Balance Measurements in Complicated Chemical Systems," *Chem. Eng. Sci.*, **24**, 947 (1969).

Vaclavek, V., and M. Loucka, "Selection of Measurements Necessary to Achieve Multicomponent Mass Balances in Chemical Plant," *Chem. Eng. Sci.*, **31**, 1199 (1976).

Vaclavek, V., M. Kubicek, and M. Loucka, "Calculation of Material Balances for Chemical Engineering Systems with Due Allowance for Measurement Errors," *Theor. Found. Chem. Eng.*, **9**, 242 (1976a) [trans. of *Teor. Osnovy Khim. Tekhnol.*, **9**, 270 (1975)].

———, "Calculation of Material Balances for Chemical Engineering Systems with Due Allowance for Errors in Measurement Classification of Stream Parameters," *Theor. Found. Chem. Eng.*, **10**, 256 (1976b) [trans. of *Teor. Osnovy Khim. Tekhnol.*, **10**, 271 (1976)].

Manuscript received Sept. 9, 1985, and revision received June 17, 1986.

See NAPS document no. 04435 for 14 pages of supplementary material. Order from NAPS c/o Microfiche Publications, P.O. Box 3513, Grand Central Station, New York, NY 10163. Remit in advance in U.S. funds only $7.75 for photocopies or $4.00 for microfiche. Outside the U.S. and Canada, add postage of $4.50 for the first 20 pages and $1.00 for each of 10 pages of material thereafter, $1.50 for microfiche postage.