ELSEVIER

Parallel Computing 27 (2001) 1117–1144

PARALLEL COMPUTING

www.elsevier.com/locate/parco

Short communication

The generalized Direction Vector I test

Weng-Long Chang, Chih-Ping Chu *

Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 701, Taiwan, ROC

Received 21 September 1998; received in revised form 3 December 2000; accepted 26 December 2000

Abstract

The Direction Vector I (DVI) test [IEEE Transaction on Parallel and Distributed Systems 4 (11) (1993) 1280] is an efficient and precise data dependence method to determine whether integer solutions exist for one-dimensional arrays with *constant* bounds under any given direction vectors. In this paper, we generalize the (DVI) test. The generalized Direction Vector I (GDVI) test can be applied towards determining whether integer solutions exist for one-dimensional arrays with both *constant* and *variable* limits under any given direction vectors, improving the precision and applicability of the DVI test. Experiments with benchmark showed that among 12 152 pairs of tested one-dimensional arrays consisting of the same pair of array references with different direction vectors, 2124 had their data dependence analysis amended by the GDVI test. That is, the GDVI test increases the success rate of the DVI test by approximately 17.5%.

Keywords: Parallelizing/Vectorizing compilers; Data dependence analysis; Parallel computing

1. Introduction

The question of whether one-dimensional array references with linear subscripts may be parallelized/vectorized depends upon the resolution of those one-dimensional array aliases. The resolution of one-dimensional array aliases is to ascertain whether two references to the same one-dimensional array within a general loop may refer to the same element of that one-dimensional array. This problem in general case can be reduced to that of checking whether a linear equation with n unknown variables has

* Corresponding author. Tel.: +886-6-27-57-575; fax: +886-6-20-84-461.
E-mail address: chucp@csie.ncku.edu.tw (C.-P. Chu).

0167-8191/01/$ - see front matter
PII: S 0 1 6 7 - 8 1 9 1 (0 1) 0 0 0 8 6 - 2

an integer solution, which satisfies the bounds for each variable in the linear equation. It is assumed that the linear equation is written as

$$a_1X_1 + a_2X_2 + \cdots + a_{n-1}X_{n-1} + a_nX_n = a_0, \tag{1.1}$$

where each a_j is an integer for $0 \leqslant j \leqslant n$. It is postulated that the limits to each variable in (1.1) are represented as

$$P_{r,0} + \sum_{s=1}^{r-1} P_{r,s}X_s \leqslant X_r \leqslant Q_{r,0} + \sum_{s=1}^{r-1} Q_{r,s}X_s, \tag{1.2}$$

where $P_{r,0}$, $Q_{r,0}$, $P_{r,s}$ and $Q_{r,s}$ are integers for $1 \leqslant r \leqslant n$.

If each of $P_{r,s}$ and $Q_{r,s}$ is zero in the constraints of (1.2), then (1.2) will be reduced to

$$P_{r,0} \leqslant X_r \leqslant Q_{r,0}, \quad \text{where } 1 \leqslant r \leqslant n. \tag{1.3}$$

That is, the bounds for each variable X_r are constants.

The Direction Vector I test proposed by Kong et al. [1] is to determine integer solutions for a linear equation (1.1) with the constraints of (1.3). In this paper, the DVI test is extended to check whether a linear equation (1.1) together with the bounds of (1.2) and any given direction vectors has a relevant integer solution. A theoretical analysis explains that we take advantage of the trapezoidal shape of the convex sets derived from a linear equation under variable limits as well as any given direction vectors in a data dependence testing. An algorithm called the generalized Direction Vector I (GDVI) test has been implemented and several measurements have also been performed. Actually, the GDVI test is equivalent to a version of the DVI test which combines the *Banerjee algorithm* [3] and the GCD test.

The rest of this paper is proffered as follows. In Section 2, the definition of data dependence is presented. The GCD test, the Banerjee test, the DVI test and the extended I test are briefly reviewed. In Section 3, the theoretical aspects and the time complexity for the GDVI test are proposed. Experimental results showing the advantages of the generalization of the DVI test are given in Section 4. Finally, in Section 5 we draw brief conclusions.

2. Background

In this section, we mainly introduce the concept of data dependence and cite some dependence testing methods.

2.1. Data dependence

It is assumed that S_1 and S_2 are two statements within a general loop. The general loop is presumed to contain d common loops. Statements S_1 and S_2 are postulated to be embedded in $d + p$ loops and $d + q$ loops, respectively. Each iteration of a general loop is identified by an iteration vector whose elements are the values of the iteration variables for that iteration. For example, the instance of the statement S_1 during iteration $\vec{i} = (i_1, \ldots, i_d, \ldots, i_{d+p})$ is denoted $S_1(\vec{i})$; the instance of the statement S_2 during iteration $\vec{j} = (j_1, \ldots, j_d, \ldots, j_{d+q})$ is denoted $S_2(\vec{j})$. If $(i_1, \ldots, i_d, \ldots, i_{d+p})$ is identical to $(j_1, \ldots, j_d, \ldots, j_{d+q})$ or $(i_1, \ldots, i_d, \ldots, i_{d+p})$ precedes $(j_1, \ldots, j_d, \ldots, j_{d+q})$ lexicographically, then $S_1(\vec{i})$ is said to precede $S_2(\vec{j})$, denoted $S_1(\vec{i}) < S_2(\vec{j})$. Otherwise, $S_2(\vec{j})$ is said to precede $S_1(\vec{i})$, denoted $S_1(\vec{i}) > S_2(\vec{j})$.

Definition 2.1. A vector of the form $\vec{\theta} = (\theta_1, \ldots, \theta_d)$ is termed as a direction vector. The direction vector $(\theta_1, \ldots, \theta_d)$ is said to be the direction vector from $S_1(\vec{i})$ to $S_2(\vec{j})$ if for $1 \leqslant k \leqslant d$, $i_k \theta_k j_k$, i.e., the relation θ_k is defined by

$$\theta_k = \begin{cases} < & \text{if } i_k < j_k, \\ = & \text{if } i_k = j_k, \\ > & \text{if } i_k > j_k, \\ * & \text{the relation of } i_k \text{ and } j_k \text{ can be ignored,} \\ & \quad \text{i.e., can be any one of } \{<, =, >\}. \end{cases}$$

Definition 2.2 [3]. Given a linear equation (1.1) beneath the constraints of (1.2) and a specific direction vector $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, where d refers to the number of common loops. If $\theta_k =$ '$=$', $1 \leqslant k \leqslant d$, $X_{2k-1} \theta_k X_{2k}$ and X_{2k-1} and X_{2k} refer to the same loop indexed variable, then the two terms, $a_{2k-1}X_{2k-1}$ and $a_{2k}X_{2k}$, in the equation will be merged and the bounds for the corresponding variable are unchanged. If $\theta_k \in \{<, >\}$, $1 \leqslant k \leqslant d$, then the bounds of (1.2) for each pair of relative variables will be redefined, assuming $X_{2k-1} \theta_k X_{2k}$ and X_{2k-1} and X_{2k} refer to the same loop indexed variable. The new constraints for X_{2k-1} and X_{2k} are either (2.1) or (2.2).

If $\theta_k = <$, then

$$\begin{aligned} & P_{2k-1,0} + \sum_{s=1}^{2k-2} P_{2k-1,s} X_s \leqslant X_{2k-1} \leqslant (Q_{2k-1,0} - 1) + \sum_{s=1}^{2k-2} Q_{2k-1,s} X_s, \\ & 1 + X_{2k-1} \leqslant X_{2k} \leqslant Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s} X_s. \end{aligned} \tag{2.1}$$

If $\theta_k = >$, then

$$\begin{aligned} & P_{2k-1,0} + \sum_{s=1}^{2k-2} P_{2k-1,s} X_s \leqslant X_{2k-1} \leqslant Q_{2k-1,0} + \sum_{s=1}^{2k-2} Q_{2k-1,s} X_s, \\ & P_{2k,0} + \sum_{s=1}^{2k-1} P_{2k,s} X_s \leqslant X_{2k} \leqslant X_{2k-1} - 1. \end{aligned} \tag{2.2}$$

2.2. The GCD, Banerjee and DVI tests

The GCD test is based upon a theorem of elementary number theory, which says that a linear equation (1.1) has an integer solution if and only if $\gcd(a_1, \ldots, a_n)$ is a divisor of a_0. The Banerjee test (involving the Banerjee inequalities and the Banerjee algorithm) computes the minimum and maximum values for the expression on the left-hand side of (1.1) beneath constant or variable constraints as well as any given direction vectors. By the Intermediate Value Theorem, the Banerjee test infers that (1.1) has a real-valued solution if and only if the minimum value is less than or equal to a_0 and the maximum value is greater than or equal to a_0.

A linear equation (1.1) with the bounds of (1.3) and any given direction vectors will be said to be integer solvable if the linear equation (1.1) has an integer solution to satisfy the bounds of (1.3) and any given direction vectors for each variable in the linear equation (1.1). The DVI test deals with a linear equation by first transforming it to an *interval* equation. Definitions 2.3 and 2.4 cited from [1] define integer intervals and an interval equation.

Definition 2.3. Let $[\alpha_1, \alpha_2]$ represent the integer intervals from α_1 to α_2, i.e., the set of all integers between α_1 and α_2.

To avoid redundancy, throughout this paper we will use the term interval to refer to the integer interval.

Definition 2.4. Let $a_1, \ldots, a_{n-1}, a_n$, L and U be integers. A linear equation

$$a_1X_1 + a_2X_2 + \cdots + a_{n-1}X_{n-1} + a_nX_n = [L, U], \tag{2.3}$$

which is referred to as an integer interval equation, will be used to denote the set of ordinary equations consisting of:

$$\begin{aligned} &a_1X_1 + a_2X_2 + \cdots + a_{n-1}X_{n-1} + a_nX_n = L, \\ &a_1X_1 + a_2X_2 + \cdots + a_{n-1}X_{n-1} + a_nX_n = L + 1, \\ &\qquad\vdots \\ &a_1X_1 + a_2X_2 + \cdots + a_{n-1}X_{n-1} + a_nX_n = U. \end{aligned}$$

Similarly, we will use the term interval equation to refer to integer interval equation throughout this paper.

An interval equation (2.3) will be said to be integer solvable if one of the equations in the set, which it defines, is integer solvable. The immediate way to determine this is to test if an integer in between L and U is divisible by the GCD of the coefficients of the left-hand-side terms. If $L > U$ in an interval equation (2.3), then there are no integer solutions for this interval equation. If the expression on the left-hand side of an interval equation (2.3) is reduced to zero items, in the processing of testing, then the interval equation (2.3) will be said to be integer solvable if and only if $L \leqslant 0 \leqslant U$. It is easy to see that the linear equation (1.1) is integer solvable if and only if the interval equation, $a_1X_1 + a_2X_2 + \cdots + a_{n-1}X_{n-1} + a_nX_n = [a_0, a_0]$, is integer solvable.

Definition 2.5 [1]. Let a variable a_i be an integer $1 \leqslant i \leqslant n$. The positive part a_i^+ and the negative part a_i^- of an integer a_i are defined by $a_i^+ = \mathrm{MAX}\{a_i, 0\}$, and $a_i^- = \mathrm{MAX}\{-a_i, 0\}$.

Theorem 2.1 [1]. *Given the interval equation* (2.3) *subject to the constraints of* (1.3) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops and for all k*, $1 \leqslant k \leqslant d$, $\theta_k = <$. *Let*

$$t = \begin{cases} \max(|a_{2k-1}|, |a_{2k}|) & \text{if } a_{2k-1} * a_{2k} > 0, \\ \max(\min(|a_{2k-1}|, |a_{2k}|), |a_{2k-1} + a_{2k}|) & \text{if } a_{2k-1} * a_{2k} < 0. \end{cases}$$

If $t \leqslant U - L + 1$, *then the interval equation*

$$a_1X_1 + \cdots + a_{2d-1}X_{2d-1} + a_{2d}X_{2d} + \cdots + a_nX_n = [L, U]$$

is ($P_{2k,0} \leqslant X_{2k-1} < X_{2k} \leqslant Q_{2k,0}$ *for* $1 \leqslant k \leqslant d$, *and* $P_{r,0} \leqslant X_r \leqslant Q_{r,0}$ *for* $2d + 1 \leqslant r \leqslant n$)*-integer solvable if and only if the interval equation*

$$\begin{aligned} a_1X_1 + \cdots + a_{2k-2}X_{2k-2} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n \\ = \big[L - (a_{2k-1}^{+} + a_{2k})^{+}(Q_{2k,0} - P_{2k,0} - 1) - (a_{2k-1} + a_{2k}) * P_{2k,0} - a_{2k}, \\ U + (a_{2k-1}^{-} - a_{2k})^{+}(Q_{2k,0} - P_{2k,0} - 1) - (a_{2k-1} + a_{2k}) * P_{2k,0} - a_{2k}\big] \end{aligned}$$

is ($P_{2p,0} \leqslant X_{2p-1} < X_{2p} \leqslant Q_{2p,0}$ *for* $1 \leqslant p \leqslant d, p \neq k$ *and* $P_{r,0} \leqslant X_r \leqslant Q_{r,0}, 2d + 1 \leqslant r \leqslant n$)*-integer solvable.*

It is very obvious from Theorem 2.1 that the DVI test considers a pair of same index variables to justify the movement of the two variables to the right. It is indicated from Theorem 2.1 that a pair of same index variables in Eq. (2.3) can be moved to the right if the coefficients of the two variables have small enough values to justify the movement of the two variables to the right. If all coefficients for variables in Eq. (2.3) have no sufficiently small values to justify the movements of variables to the right, then Theorem 2.1 cannot be applied. While every variable in Eq. (2.3) cannot be moved to the right, Theorem 2.2 describes a transformation using the GCD test which enables additional variables to be moved.

Theorem 2.2 [1]. *Given the interval equation* (2.3) *subject to the constraints of* (1.3) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops and for all k*, $1 \leqslant k \leqslant d$, $\theta_k = <$. *Let* $g = \gcd(a_1, \ldots, a_{n-1}, a_n)$. *The interval equation*

$$a_1X_1 + \cdots + a_{2d}X_{2d} + \cdots + a_nX_n = [L, U]$$

is ($P_{2k,0} \leqslant X_{2k-1} < X_{2k} \leqslant Q_{2k,0}$ *for* $1 \leqslant k \leqslant d$, *and* $P_{r,0} \leqslant X_r \leqslant Q_{r,0}$ *for* $2d + 1 \leqslant r \leqslant n$)*-integer solvable if and only if the interval equation*

$$(a_1/g)X_1 + \cdots + (a_{2d}/g)X_{2d} + \cdots + a_nX_n = [\lceil L/g \rceil, \lfloor U/g \rfloor]$$

is ($P_{2k,0} \leqslant X_{2k-1} < X_{2k} \leqslant Q_{2k,0}$ *for* $1 \leqslant k \leqslant d$, *and* $P_{r,0} \leqslant X_r \leqslant Q_{r,0}$ *for* $2d + 1 \leqslant r \leqslant n$)*-integer solvable.*

2.3. The extended I test

The I test [15] is an efficient and precise data dependence method to ascertain whether integer solutions exist for one-dimensional arrays with *constant* bounds.

For one-dimensional arrays with variable limits, the I test assumes that there may exist integer solutions. In our previous work [11], we extended the I test. The extended I test can be applied towards determining whether integer solutions exist for one-dimensional arrays with either *variable* or *constant* limits, improving the applicability of the I test.

3. The generalization of the DVI test

A data dependence problem is considered where one-dimensional array references are linear in terms of index variables. Bounds for one-dimensional array references are presumed to be variable. (Note: constant limits are simply a special case of variable constraints, so variable bounds actually cover constant limits.) Given the data dependence problem as specified, the generalized version of the DVI test examines a linear equation (1.1) with the constraints of (1.2) as well as any given direction vectors and deduces whether the system has integer solutions.

3.1. Basic concepts

A linear equation (1.1) under the constraints of (1.2) as well as any given direction vectors will be said to be integer solvable if the linear equation has an integer solution to satisfy the constraints of (1.2) and any given direction vectors for each variable in the linear equation.

Definition 3.1. Suppose that the constraints of X_r for $1 \leqslant r \leqslant n$ are equal to the bounds of (1.2). Let $[b_0 + \sum_{r=1}^{n} b_r X_r,\ c_0 + \sum_{r=1}^{n} c_r X_r]$ denote a variable interval in contrast to the (constant) integer interval defined in Definition 2.3, representing the set of all the integer intervals obtained by replacing X_r with values within their bounds of (1.2), where b_0, c_0, b_r and c_r for $1 \leqslant r \leqslant n$ are integers. So, the variable interval Φ is equal to

$$\left\{ \left[b_0 + \sum_{r=1}^{n} b_r x_r, c_0 + \sum_{r=1}^{n} c_r x_r \right] \middle| P_{r,0} + \sum_{s=1}^{r-1} P_{r,s} x_s \leqslant x_r \leqslant Q_{r,0} + \sum_{s=1}^{r-1} Q_{r,s} x_s \text{ for } 1 \leqslant r \leqslant n \right\}.$$

Definition 3.2a. Suppose that the constraints of X_r for $1 \leqslant r \leqslant n$ are equal to the bounds of (1.2). Let $L = b_0 + \sum_{r=1}^{n} b_r X_r$ and $U = c_0 + \sum_{r=1}^{n} c_r X_r$, where $L \leqslant U$, and b_0, c_0, b_r and c_r for $1 \leqslant r \leqslant n$ are integers. Let $a_1, \ldots, a_{n-1}$ and a_n be integers. The following equation

$$a_1 X_1 + a_2 X_2 + \cdots + a_{n-1} X_{n-1} + a_n X_n = [L, U] \tag{3.1}$$

is referred as a *variable interval equation*, in contrast to the (constant) *interval equation* defined in Definition 2.4, which will be used to denote the set of all the interval equations inferred from every variable X_r within the bounds of (1.2). The variable interval equation Ψ is denoted to be equal to the set of all the (constant) interval equations; that is

$$\Psi = \left\{ a_1x_1 + \cdots + a_nx_n = \left[b_0 + \sum_{r=1}^{n} b_rx_r, c_0 + \sum_{r=1}^{n} c_rx_r \right] \middle| P_{r,0} + \sum_{s=1}^{r-1} P_{r,s}x_s \leqslant x_r \leqslant Q_{r,0} + \sum_{s=1}^{r-1} Q_{r,s}x_s \text{ for } 1 \leqslant r \leqslant n \right\}.$$

If $b_0 + \sum_{r=1}^{n} b_rx_r \leqslant \sum_{r=1}^{n} a_rx_r \leqslant c_0 + \sum_{r=1}^{n} c_rx_r$ in the variable interval equation (3.1), then the corresponding constant interval equation exists. Otherwise, the corresponding one does not exist. If each of b_r and c_r in the variable interval equation (3.1) is zero for $1 \leqslant r \leqslant n$, then the variable interval equation only contains one interval equation.

Definition 3.2b. The variable interval equation (3.1) will be said to be $(P_1, Q_1; \ldots; P_n, Q_n)$-integer solvable if one of the (constant) interval equations in the set Ψ is $(P_1, Q_1; \ldots; P_n, Q_n)$-integer solvable, where P_k, Q_k, $1 \leqslant k \leqslant n$, in $(P_1, Q_1; \ldots; P_n, Q_n)$ refer to the lower and upper bounds of the variable X_k defined in (1.2).

If $b_0 + \sum_{r=1}^{n} b_rx_r > c_0 + \sum_{r=1}^{n} c_rx_r$ in every interval equation in the set Ψ, then there are no integer solutions for the variable interval equation. If $\sum_{r=1}^{n} a_rx_r < b_0 + \sum_{r=1}^{n} b_rx_r$ or $\sum_{r=1}^{n} a_rx_r > c_0 + \sum_{r=1}^{n} c_rx_r$ in every interval equation in the set Ψ, then there are no integer solutions for the set Ψ. If all of the integers in between the intervals of the interval equations in the set Ψ are not divisible by the greatest common divisor of the left-hand-side coefficients of (3.1), then the variable interval equation will be integer unsolvable. If the expression of the left-hand side for one of the interval equations in the set Ψ becomes zero items, in the processing of testing, then the variable interval equation will be said to be integer solvable if and only if $b_0 + \sum_{r=1}^{n} b_rx_r \leqslant 0 \leqslant c_0 + \sum_{r=1}^{n} c_rx_r$.

Definition 3.2c. Given a variable interval equation (3.1) subject to the constraints of (1.2) and a specific direction vector $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, where d is the number of common loops. This variable interval equation will be said to be $(P_1, Q_1; \ldots; P_n, Q_n)$-integer solvable if one of the (constant) interval equations in the set Ψ is $(P_1, Q_1; \ldots; P_n, Q_n)$-integer solvable, where $P_k, Q_k, 1 \leqslant k \leqslant n$, in $(P_1, Q_1; \ldots; P_n, Q_n)$ refer to the lower and upper bounds of the variable X_k defined in (1.2) but may be redefined due to dependence direction θ_k (according to Definition 2.2).

For example, suppose the variable interval equation (3.1) under the constraints (1.2) and a specific direction vector $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, where d is the number of common loops, θ_k, $1 \leqslant k \leqslant d$, is equal to "$<$", θ_p, for all $1 \leqslant p \leqslant d$ and $p \neq k$, are equal to "$*$". This variable interval equation is (Rel. 3.2)-integer solvable if one of the interval equation in the set Ψ is (Rel. 3.2)-integer solvable, where (Rel. 3.2) is

$$\Bigg(P_{2k-1,0}+\sum_{s=1}^{2k-2}P_{2k-1,s}X_s \leqslant X_{2k-1} \leqslant (Q_{2k-1,0}-1)+\sum_{s=1}^{2k-2}Q_{2k-1,s}X_s \text{ and}$$

$$1+X_{2k-1} \leqslant X_{2k} \leqslant Q_{2k,0}+\sum_{s=1}^{2k-1}Q_{2k,s}X_s,\ 1\leqslant k\leqslant d;\ P' \leqslant X_{2p-1} \leqslant Q' \text{ and}$$

$$P'' \leqslant X_{2p} \leqslant Q'',\ \text{for } 1\leqslant p\leqslant d, p\neq k,\ \text{the values of } P', Q', P'', Q''$$

$$\text{will depend on the dependence direction } \theta_p;\ P_{r,0}+\sum_{s=1}^{r-1}P_{r,s}X_s \leqslant X_r \leqslant Q_{r,0}$$

$$+\sum_{s=1}^{r-1}Q_{r,s}X_s \text{ for } 2d+1\leqslant r\leqslant n\Bigg). \qquad \text{(Rel.3.2)}$$

Definition 3.3. Suppose that the constraints of X_r for $1\leqslant r\leqslant n$ are equal to the bounds of (1.2). Let K represent the left-hand-side expression $\sum_{r=1}^{n} a_r X_r$, $L=b_0+\sum_{r=1}^{n} b_r X_r$ and $U=c_0+\sum_{r=1}^{n} c_r X_r$ in Eq. (3.1), where $L\leqslant K\leqslant U$, and b_0, c_0, a_r, b_r and c_r for $1\leqslant r\leqslant n$ are integers. The set of length of the right-hand-side variable interval, Ω, in the variable interval equation (3.1) is denoted to be equal to

$$\left\{1+(c_0-b_0)+\sum_{r=1}^{n}(c_r-b_r)x_r \,\middle|\, P_{r,0}+\sum_{s=1}^{r-1}P_{r,s}x_s \leqslant x_r \leqslant Q_{r,0}\right.$$

$$\left.+\sum_{s=1}^{r-1}Q_{r,s}x_s \text{ for } 1\leqslant r\leqslant n\right\}.$$

It is easy to see that a linear equation (1.1) is integer solvable if and only if the only interval equation

$$a_1X_1+a_2X_2+\cdots+a_{n-1}X_{n-1}+a_nX_n=[a_0,a_0]$$

in the set Ψ is integer solvable. The following theorem shows the way to obtain the maximum length of the variable interval in a variable interval equation.

Theorem 3.1 [1]. *The maximum element in the set Ω denoted in Definition* 3.3 *can be determined by the Banerjee algorithm.*

The GDVI test involves a number of variable-interval-equation to variable-interval-equation transformations. In Sections 3.2–3.5, Theorems 3.2–3.7, and Lemmas 3.2–3.8 are extensions of Theorems 2.1 and 2.2. These theorems and lemmas will be employed towards doing transformations of variable interval equation.

3.2. Transformation of variable interval equation with "<" in direction vector

The bounds to each pair of same index variables in the variable interval equation (3.1) will be redefined to be the constraints of (2.1) denoted in Definition 2.2 if the direction vector "<" are considered. Theorems 3.2 and 3.3 and Lemmas 3.2 and 3.3 are derived to describe the way of variable-interval-equation to variable-interval-equation transformations under variable limits as well as a specific direction vector $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, where d is the number of common loops, θ_k is equal to "<", $1 \leqslant k \leqslant d$, and θ_p are equal to "$*$" for $1 \leqslant p \leqslant d$, $p \neq k$.

Theorem 3.2. *Given a variable interval equation* (3.1), *defined in Definition* 3.2a, *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where* d *is the number of common loops*, θ_k *is equal to* "<", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d, p \neq k$. *If* $a_{2k} > 0, a_{2k} \geqslant b_{2k} \geqslant 0, a_{2k} \geqslant c_{2k} \geqslant 0$, *and the value for* a_{2k} *is less than or equal to the length of the right-hand-side interval of one of the interval equations in the variable interval equation*, *then the variable interval equation is* (Rel. 3.2)-*integer solvable if and only if the variable interval equation*

$$\begin{aligned} &a_1X_1 + \cdots + a_{2k-1}X_{2k-1} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n \\ &= \left[b_0 + \sum_{r=1}^{n} b_rX_r + (b_{2k} - a_{2k})\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}X_s \right), \right. \\ &\qquad \left. c_0 + \sum_{r=1}^{n} c_rX_r + (c_{2k} - a_{2k})(1 + X_{2k-1}) \right] \quad \textit{for } r \neq 2k \end{aligned} \tag{3.2}$$

is (*Rel.* 3.3)-*integer solvable*, *where* (*Rel.* 3.3) *is*

$$\left(\begin{aligned} &P' \leqslant X_{2p-1} \leqslant Q' \textit{ and } P'' \leqslant X_{2p} \leqslant Q'' \textit{ for } 1 \leqslant p \leqslant d \textit{ and } p \neq k, \\ &\textit{the values of } P', Q', P'' \textit{ and } Q'' \textit{ will depend on the direction } \theta_p; \\ &P_{2k-1,0} + \sum_{s=1}^{2k-2} P_{2k-1,s}X_s \leqslant X_{2k-1} \leqslant (Q_{2k-1,0} - 1) + \sum_{s=1}^{2k-2} Q_{2k-1,s}X_s; \textit{ and} \\ &P_{r,0} + \sum_{s=1}^{r-1} P_{r,s}X_s \leqslant X_r \leqslant Q_{r,0} + \sum_{s=1}^{r-1} Q_{r,s}X_s \textit{ for } 2d+1 \leqslant r \leqslant n \end{aligned} \right). \tag{Rel.3.3}$$

Proof. (*Only if*) According to Definition 3.2a, the variable interval equation (3.1) represents the set of all the interval equations inferred from every variable X_r with the bounds of (1.2) and a specific direction vector $\vec{\theta} = (\theta_1, \ldots, \theta_d)$. The set of all the interval equations Ψ is equal to

$$\left\{ a_1x_1 + \cdots + a_nx_n = \left[b_0 + \sum_{r=1}^{n} b_rx_r, c_0 + \sum_{r=1}^{n} c_rx_r \right] \middle| \text{ for each } x_r, 1 \leqslant r \leqslant n, \right. \\ \left. \text{satisfies its bounds} \right\}.$$

Since the set of all the interval equations Ψ is integer solvable, so there exists an interval equation in the set Ψ that is integer solvable. Let the interval equation be

$$a_1x_1 + \cdots + a_{n-1}x_{n-1} + a_nx_n = \left[b_0 + \sum_{r=1}^{n} b_rx_r, c_0 + \sum_{r=1}^{n} c_rx_r\right],$$

where $x_1, \ldots, x_n$ satisfy the conditions for (Rel. 3.2)-integer solvable.

We thus have that

$$a_1x_1 + \cdots + a_nx_n = \left[(b_{2k} - a_{2k})x_{2k} + b_0 + \sum_{r=1}^{n} b_rx_r, (c_{2k} - a_{2k})x_{2k} + c_0 + \sum_{r=1}^{n} c_rx_r\right]$$

for $r \neq 2k$.

Let the integer interval Λ be

$$\left[(b_{2k} - a_{2k})x_{2k} + b_0 + \sum_{r=1}^{n} b_rx_r, (c_{2k} - a_{2k})x_{2k} + c_0 + \sum_{r=1}^{n} c_rx_r\right]$$

for $r \neq 2k$, and let the integer interval Δ be

$$\left[b_0 + \sum_{r=1}^{n} b_rx_r + (b_{2k} - a_{2k})\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s\right), c_0 + \sum_{r=1}^{n} c_rx_r + (c_{2k} - a_{2k})(1 + x_{2k-1})\right]$$

for $r \neq 2k$.

Since $0 \leqslant b_{2k} \leqslant a_{2k}$ and $0 \leqslant c_{2k} \leqslant a_{2k}$ according to the assumption of Theorem 3.2, the integer interval Λ is enclosed by the integer interval Δ. Therefore, the interval equation

$$a_1x_1 + \cdots + a_{2k-1}x_{2k-1} + a_{2k+1}x_{2k+1} + \cdots + a_nx_n = \left[b_0 + \sum_{r=1}^{n} b_rx_r + (b_{2k} - a_{2k})\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s\right), c_0 + \sum_{r=1}^{n} c_rx_r + (c_{2k} - a_{2k})(1 + x_{2k-1})\right]$$

for $r \neq 2k$ is (Rel. 3.3)-integer solvable.

The interval equation lies in the set of all the interval equations concluded from the variable interval equation (3.2), so the variable interval equation (3.2) is (Rel. 3.3)-integer solvable.

(*If*): According to Definition 3.2a, the variable interval equation (3.2) represents the set of all the interval equations inferred from every variable X_r with the bounds of (1.2) and a specific direction vector $\vec{\theta} = (\theta_1, \ldots, \theta_d)$. The set of all the interval equations Ψ is equal to

$$\Big\{a_1x_1 + \cdots + a_{2k-1}x_{2k-1} + a_{2k+1}x_{2k+1} + \cdots + a_nx_n$$

$$= \left[b_0 + \sum_{r=1}^{n} b_rx_r + (b_{2k} - a_{2k})\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s\right), c_0 + \sum_{r=1}^{n} c_rx_r\right.$$

$$\left. + (c_{2k} - a_{2k})(1 + x_{2k-1})\right]\Bigg| P' \leqslant X_{2p-1} \leqslant Q' \text{ and } P'' \leqslant X_{2p} \leqslant Q'' \text{ for } 1 \leqslant p \leqslant d$$

and $p \neq k$, the values of P', Q', P'' and Q'' will depend on the direction θ_p;

$$P_{2k-1,0} + \sum_{s=1}^{2k-2} P_{2k-1,s}x_s \leqslant x_{2k-1} \leqslant (Q_{2k-1,0} - 1) + \sum_{s=1}^{2k-2} Q_{2k-1,s}x_s, \text{ and}$$

$$P_{r,0} + \sum_{s=1}^{r-1} P_{r,s}x_s \leqslant x_r \leqslant Q_{r,0} + \sum_{s=1}^{r-1} Q_{r,s}x_s \text{ for } 2d + 1 \leqslant r \leqslant n\Big\}.$$

The set of all the interval equations Ψ is integer solvable, so there exists an interval equation that is integer solvable in the set Ψ. Let the interval equation be

$$\begin{aligned} &a_1x_1 + \cdots + a_{2k-1}x_{2k-1} + a_{2k+1}x_{2k+1} + \cdots + a_nx_n \\ &\quad = \left[b_0 + \sum_{r=1}^{n} b_rx_r + (b_{2k} - a_{2k})\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s\right),\right. \\ &\qquad \left. c_0 + \sum_{r=1}^{n} c_rx_r + (c_{2k} - a_{2k})(1 + x_{2k-1})\right] \end{aligned}$$

for $r \neq 2k$, where $x_1, \ldots, x_{2k-1}, x_{2k+1}, \ldots, x_n$ satisfy the conditions for (Rel. 3.3)-integer solvable. Since $0 \leqslant b_{2k} \leqslant a_{2k}$ and $0 \leqslant c_{2k} \leqslant a_{2k}$ according to the assumption of Theorem 3.2 and suppose $1 + x_{2k-1} \leqslant x_{2k} \leqslant Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s$, we obtain the following results:

$$\begin{aligned} &b_0 + \sum_{r=1}^{n} b_rx_r + (b_{2k} - a_{2k})\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s\right) \\ &\quad \geqslant b_0 + \sum_{r=1}^{n} b_rx_r + b_{2k}x_{2k} - a_{2k}\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s\right) \quad \text{for } r \neq 2k, \end{aligned}$$

and

$$\begin{aligned} &c_0 + \sum_{r=1}^{n} c_rx_r + (c_{2k} - a_{2k})(1 + x_{2k-1}) \\ &\quad \leqslant c_0 + \sum_{r=1}^{n} c_rx_r + c_{2k}x_{2k} - a_{2k}(1 + x_{2k-1}) \quad \text{for } r \neq 2k. \end{aligned}$$

Hence, the interval equation

$$a_1x_1 + \cdots + a_{2k-1}x_{2k-1} + a_{2k+1}x_{2k+1} + \cdots + a_nx_n = \left[b_0 + \sum_{r=1}^{n} b_rx_r - a_{2k}\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s\right), c_0 + \sum_{r=1}^{n} c_rx_r - a_{2k}(1 + x_{2k-1})\right]$$

is (Rel. 3.2)-integer solvable.

Because the interval equation is integer solvable, a linear equation in the interval equation is integer solvable. Let the linear equation be equal to

$$a_1x_1 + \cdots + a_{2k-1}x_{2k-1} + a_{2k+1}x_{2k+1} + \cdots + a_nx_n = c,$$

where

$$b_0 + \sum_{r=1}^{n} b_rx_r - a_{2k}\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s\right) \leqslant c \leqslant c_0 + \sum_{r=1}^{n} c_rx_r - a_{2k}(1 + x_{2k-1}).$$

Let L^1 and U^1 represent $b_0 + \sum_{r=1}^{n} b_rx_r$ and $c_0 + \sum_{r=1}^{n} c_rx_r$, and let α and β represent $1 + x_{2k-1}$ and $Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s$. Consider the set of integer intervals $\{[L^1 - a_{2k}(\beta - m),\ U^1 - a_{2k}(\beta - m)] \mid 0 \leqslant m \leqslant \beta - \alpha\}$. Since $a_{2k} > 0$, these integer intervals lie in the following sequence in order of initial element:

$$\begin{array}{l}[L^1 - a_{2k}\beta,\ U^1 - a_{2k}\beta]\\ [L^1 - a_{2k}(\beta - 1),\ U^1 - a_{2k}(\beta - 1)]\\ \qquad\qquad\vdots\\ [L^1 - a_{2k}\alpha,\ U^1 - a_{2k}\alpha].\end{array}$$

The length of each integer interval is $U^1 - L^1 + 1$. Consider two consecutive integer intervals $[L^1 - a_{2k}\beta,\ U^1 - a_{2k}\beta]$ and $[L^1 - a_{2k}(\beta - 1),\ U^1 - a_{2k}(\beta - 1)]$. There is a gap between the two integer intervals if and only if $U^1 - a_{2k}\beta + 1 < L^1 - a_{2k}(\beta - 1)$ which reduces to $a_{2k} > U^1 - L^1 + 1$. But $a_{2k} \leqslant U^1 - L^1 + 1$ due to the assumption of Theorem 3.2, the gap between the two integer intervals does not exist. We thus have that

$$\left[b_0 + \sum_{r=1}^{n} b_rx_r - a_{2k}\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s\right), c_0 + \sum_{r=1}^{n} c_rx_r - a_{2k}(1 + x_{2k-1})\right]$$
$$= \left\{\left[b_0 + \sum_{r=1}^{n} b_rx_r - a_{2k}\left(Q_{2k,0} + \left(\sum_{s=1}^{2k-1} Q_{2k,s}x_s\right) - m\right),\right.\right.$$
$$\left.\left. c_0 + \sum_{r=1}^{n} c_rx_r - a_{2k}\left(Q_{2k,0} + \left(\sum_{s=1}^{2k-1} Q_{2k,s}x_s\right) - m\right)\right] \middle| 0 \leqslant m \leqslant Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s - (1 + x_{2k-1})\right\}.$$

Let $0 \leqslant p \leqslant c_0 - b_0 + \sum_{r=1}^{n} (c_r - b_r)x_r$ be the specific integer such that

$$c = b_0 + \sum_{r=1}^{n} b_rx_r - a_{2k}\left(Q_{2k,0} + \left(\sum_{s=1}^{2k-1} Q_{2k,s}x_s\right) - m\right) + p.$$

We thus have from $a_1x_1 + \cdots + a_{2k-1}x_{2k-1} + a_{2k+1}x_{2k+1} + \cdots + a_nx_n = c$ that

$$a_1x_1 + \cdots + a_nx_n = b_0 + \sum_{r=1}^{n} b_rx_r - a_{2k}\left(Q_{2k,0} + \left(\sum_{s=1}^{2k-1} Q_{2k,s}x_s\right) - m\right) + p.$$

This makes it clear that

$$a_1x_1 + \cdots + a_{2k-1}x_{2k-1} + a_{2k}\left(Q_{2k,0} + \left(\sum_{s=1}^{2k-1} Q_{2k,s}x_s\right) - m\right) + \cdots + a_nx_n$$
$$= b_0 + \sum_{r=1}^{n} b_rx_r + p.$$

But,

$$1 + x_{2k-1} \leqslant Q_{2k,0} + \left(\sum_{s=1}^{2k-1} Q_{2k,s}x_s\right) - m \leqslant Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}x_s,$$

and

$$b_0 + \sum_{r=1}^{n} b_rx_r \leqslant b_0 + \sum_{r=1}^{n} b_rx_r + p \leqslant c_0 + \sum_{r=1}^{n} c_rx_r,$$

which imply that the interval equation

$$a_1x_1 + \cdots + a_nx_n = \left[b_0 + \sum_{r=1}^{n} b_rx_r, c_0 + \sum_{r=1}^{n} c_rx_r\right]$$

is (Rel. 3.2)-integer solvable.

The interval equation lies in the set of all the interval equations concluded from the variable interval equation (3.1), so the variable interval equation (3.1) is (Rel. 3.2)-integer solvable. □

Lemma 3.2. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* "<", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d$, $p \neq k$. *If* $a_{2k} > 0$, $a_{2k} \geqslant b_{2k} \geqslant 0$, $a_{2k} \geqslant c_{2k} \geqslant 0$, *and the value for* a_{2k} *is less than or equal to the maximum length of the right-hand-side intervals of the interval equations in the variable interval equation, then the variable interval equation is* (Rel. 3.2)-*integer solvable if and only if the variable interval equation*

$$a_1X_1 + \cdots + a_{2k-1}X_{2k-1} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n$$
$$= \left[b_0 + \sum_{r=1}^{n} b_rx_r + (b_{2k} - a_{2k})\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s}X_s\right),\right.$$
$$\left. c_0 + \sum_{r=1}^{n} c_rx_r + (c_{2k} - a_{2k})(1 + X_{2k-1})\right] \quad \textit{for } r \neq 2k \tag{3.3}$$

is (Rel. 3.3)-*integer solvable.*

Proof. Refer to Theorems 3.1 and 3.2. □

Theorem 3.3. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* "<", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d$, $p \neq k$. *If* $a_{2k} < 0$, $a_{2k} \leqslant b_{2k} \leqslant 0$, $a_{2k} \leqslant c_{2k} \leqslant 0$, *and the negative value for* a_{2k} *is less than or equal to the length of the right-hand-side interval of one of the interval equations in the variable interval equation, then the variable interval equation is* (Rel. 3.2)-*integer solvable if and only if the variable interval equation*

$$a_1X_1 + \cdots + a_{2k-1}X_{2k-1} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n$$
$$= \left[b_0 + \sum_{r=1}^{n} b_r x_r + (b_{2k} - a_{2k})(1 + X_{2k-1}), \right.$$
$$\left. c_0 + \sum_{r=1}^{n} c_r x_r + (c_{2k} - a_{2k})\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s} X_s\right)\right] \quad \textit{for } r \neq 2k \tag{Rel.3.4}$$

is (Rel. 3.3)-*integer solvable.*

Proof. Similar to Theorem 3.2. □

Lemma 3.3. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* "<", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d$, $p \neq k$. *If* $a_{2k} < 0$, $a_{2k} \leqslant b_{2k} \leqslant 0$, $a_{2k} \leqslant c_{2k} \leqslant 0$, *and the negative value for* a_{2k} *is less than or equal to the maximum length of the right-hand-side intervals of the interval equations in the variable interval equation, then the variable interval equation is* (Rel. 3.2)-*integer solvable if and only if the variable interval equation*

$$a_1X_1 + \cdots + a_{2k-1}X_{2k-1} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n$$
$$= \left[b_0 + \sum_{r=1}^{n} b_r x_r + (b_{2k} - a_{2k})(1 + X_{2k-1}), \right.$$
$$\left. c_0 + \sum_{r=1}^{n} c_r x_r + (c_{2k} - a_{2k})\left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s} X_s\right)\right] \quad \textit{for } r \neq 2k \tag{Rel.3.5}$$

is (Rel. 3.3)-*integer solvable.*

Proof. Refer to Theorems 3.1 and 3.3. □

The set of all the interval equations Ψ at least contains one interval equation in light of Definition 3.2a. The set Ω for the length of the right-hand-side interval on every interval equation in the set Ψ at least consists of one element according to Definitions 3.2a and 3.3. If the set Ω only includes one element, then the only element will be equal to $c_0 - b_0 + 1$ according to Definition 3.3. Otherwise, every element in the set will be equal to $1 + (c_0 - b_0) + \sum_{r=1}^{n}(c_r - b_r)x_r$. It is obvious from Definitions 3.1,3.2a–3.2c and 3.3 that the value for every element in the set Ω is greater than or equal to one. Theorems 3.2 and 3.3 and Lemmas 3.2 and 3.3 can be used to determine integer solutions for the variable interval equation (3.1) with the bounds of (1.2) and with "$<$" in direction vectors if the coefficient a_{2k} for one item in Eq. (3.1) is small enough to justify the movement of the item to the right. If $a_{2k} > 0$, $a_{2k} \geqslant b_{2k} \geqslant 0$, $a_{2k} \geqslant c_{2k} \geqslant 0$, and the value for a_{2k} is less than or equal to one of the elements in the set Ω in light of the assumption of Theorem 3.2 and Lemma 3.2, then the value is actually equivalent to the small enough value to justify the movement of the item to the right. If $a_{2k} < 0$, $a_{2k} \leqslant b_{2k} \leqslant 0$, $a_{2k} \leqslant c_{2k} \leqslant 0$, and the negative value for a_{2k} is less than or equal to one of the elements in the set Ω in light of the assumption of Theorem 3.3 and Lemma 3.3, then the value is actually equivalent to the small enough value to justify the movement of the item to the right. On the other hand, Theorems 3.2 and 3.3 and Lemmas 3.2 and 3.3 are inapplicable towards ascertaining integer solutions of the variable interval equation (3.1) if the absolute values of the coefficients for all the items in the variable interval equation (3.1) are greater than the maximum element in the set Ω. As proved in Theorem 3.1, the Banerjee algorithm can be employed to determine the maximum element in the set Ω.

We now give an example to show the precision of the GDVI test is over that of the DVI test, when it is applied to deal with a data dependence problem for a linear equation with *constant* bounds under a given direction vector.

Consider the equation

$$-3X_1 + X_2 = 10 \tag{Ex1}$$

subject to the bounds

$$1 \leqslant X_1 \leqslant 100 \quad \text{and} \quad 1 \leqslant X_2 \leqslant 100,$$

and the limits of a direction vector

$$X_1 < X_2.$$

The linear equations (Ex1) are rewritten as the interval equation

$$-3X_1 + X_2 = [10, 10]. \tag{Ex1.1}$$

If Theorem 2.1 is used to resolve the interval equation (Ex1.1), then a value for t is gained to be equal to two. The two terms X_1 and X_2 cannot be moved to the right-hand side of (Ex1.1) because there is no suitable value for t to justify the movement of the two terms X_1 and X_2 to the right-hand side of (Ex1.1). Therefore, the DVI test *assumes* that there are integer solutions.

According to the constraints of (2.1) denoted in Definition 2.2, the constraints for X_1 and X_2 will be redefined by $1 \leqslant X_1 \leqslant 99$, and $1 + X_1 \leqslant X_2 \leqslant 100$. If the GDVI test is used to resolve the same problem, then in light of Definitions 3.2a–3.2c the set of all the interval equations Ψ is equal to

$$\{-3X_1 + X_2 = [10, 10] \mid 1 \leqslant X_1 \leqslant 99 \text{ and } 1 + X_1 \leqslant X_2 \leqslant 100\}.$$

The set Ω for the length of the right-hand-side interval on every interval equation in the set Ψ is $\{1\}$. Therefore, the maximum element in the set Ω is one. It is obvious from Definitions 3.2a–3.2c that the set Ψ is integer solvable if the only interval equation in the set Ψ is integer solvable. The coefficient for X_2 satisfies the assumption of Lemma 3.2: (1) $1 > 0$, (2) $1 \geqslant 0 \geqslant 0$, (3) $1 \geqslant 0 \geqslant 0$ and (4) the value of the coefficient is equal to one. Lemma 3.2 is applied towards moving the term X_2 to the right-hand side of the only interval equation in the set Ψ. The variable interval equation (i.e., the new set Ψ_1) in light of Lemma 3.2 and Definitions 3.2a–3.2c is

$$\{-3X_1 = [-90, 9 - X_1] \mid 1 \leqslant X_1 \leqslant 99\}.$$

Now the set Ω_1 for the length of the right-hand-side interval of the interval equations in the set Ψ_1 is equal to $\{-X_1 + 100 \mid 1 \leqslant X_1 \leqslant 99\}$. The maximum element computed by the Banerjee algorithm in the set Ω_1 is 99. When the maximum element is 99, the value for X_1 is equal to 1. The value to X_1 is one, so $-3 = [-90, 8]$ $(-90 \leqslant -3 \leqslant 8)$ hold. Therefore, there exists a constant interval equation in the set Ψ_1 satisfying the given limitations. The coefficient for X_1 satisfies the assumption of Lemma 3.3: (1) $-3 < 0$, (2) $-3 \leqslant 0 \leqslant 0$, (3) $-3 \leqslant -1 \leqslant 0$, and (4) the negative value of the coefficient is less than 99. Lemma 3.3 is employed toward moving the term $-3X_1$ to the right. The new set Ψ_2 is

$$\{0 = [-87, 207]\}.$$

The expression of the left-hand side of the only interval equation in the set Ψ_2 is reduced to zero items. The only interval equation in the set Ψ_2 is integer solvable because $-87 \leqslant 0 \leqslant 207$ is true. Therefore, the GDVI test *concludes* that there are integer solutions.

3.3. Transformation of variable interval equation with ">" in direction vector

The bounds to each pair of same index variables in the variable interval equation (3.1) will be redefined to be the constraints of (2.2) denoted in Definition 2.2 when the direction vector ">" are considered. Theorems 3.4 and 3.5 and Lemmas 3.4 and 3.5 are derived to describe the way of variable-interval-equation to variable-interval-equation transformations under variable limits as well as a specific direction vector $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, where d is the number of common loops, θ_k is equal to ">", $1 \leqslant k \leqslant d$, and θ_p are equal to "$*$" for $1 \leqslant p \leqslant d$, $p \neq k$.

Theorem 3.4. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* ">", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d$, $p \neq k$. *If* $a_{2k} > 0$, $a_{2k} \geqslant b_{2k} \geqslant 0$, $a_{2k} \geqslant c_{2k} \geqslant 0$, *and the value for* a_{2k} *is less than or equal to the length of the right-hand-side interval on one of the interval equations in the variable interval equation*, *then the variable interval equation is* (Rel. 3.4)-*integer solvable*, *where* (Rel. 3.4) *is*

$$\Bigg(P_{2k-1,0} + \sum_{s=1}^{2k-2} P_{2k-1,s}X_s \leqslant X_{2k-1} \leqslant Q_{2k-1,0} + \sum_{s=1}^{2k-2} Q_{2k-1,s}X_s, \text{ and}$$

$$P_{2k,0} + \sum_{s=1}^{2k-1} P_{2k,s}X_s \leqslant X_{2k} \leqslant X_{2k-1} - 1, \text{ where } 1 \leqslant k \leqslant d;\ P' \leqslant X_{2p-1} \leqslant Q' \text{ and}$$

$$P'' \leqslant X_{2p} \leqslant Q'' \text{ for } 1 \leqslant p \leqslant d,\ p \neq k, \text{ the values of } P', Q', P'', Q''$$

$$\text{will depend on the dependence direction } \theta_p;$$

$$P_{r,0} + \sum_{s=1}^{r-1} P_{r,s}X_s \leqslant X_r \leqslant Q_{r,0} + \sum_{s=1}^{r-1} Q_{r,s}X_s \text{ for } 2d+1 \leqslant r \leqslant n\Bigg), \quad \text{(Rel.3.4)}$$

if and only if the variable interval equation

$$a_1X_1 + \cdots + a_{2k-1}X_{2k-1} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n$$
$$= \Bigg[b_0 + \sum_{r=1}^{n} b_r x_r + (b_{2k} - a_{2k})(X_{2k-1} - 1),$$
$$c_0 + \sum_{r=1}^{n} c_r x_r + (c_{2k} - a_{2k})\left(P_{2k,0} + \sum_{s=1}^{2k-1} P_{2k,s}X_s\right)\Bigg] \quad \text{for } r \neq 2k \qquad (3.6)$$

is (Rel. 3.5)-*integer solvable*, *where* (Rel. 3.5) *is*

$$\Bigg(P' \leqslant X_{2p-1} \leqslant Q' \text{ and } P'' \leqslant X_{2p} \leqslant Q'' \text{ for } 1 \leqslant p \leqslant d,\ p \neq k,$$

$$\text{the values of } P', Q', P'', Q'' \text{ will depend on the dependence direction } \theta_p;$$

$$P_{2k-1,0} + \sum_{s=1}^{2k-2} P_{2k-1,s}X_s \leqslant X_{2k-1} \leqslant Q_{2k-1,0} + \sum_{s=1}^{2k-2} Q_{2k-1,s}X_s, \text{ and}$$

$$P_{r,0} + \sum_{s=1}^{r-1} P_{r,s}X_s \leqslant X_r \leqslant Q_{r,0} + \sum_{s=1}^{r-1} Q_{r,s}X_s \text{ for } 2d+1 \leqslant r \leqslant n\Bigg). \quad \text{(Rel.3.5)}$$

Proof. Refer to Theorem 3.2. □

Lemma 3.4. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops*, θ_k *is equal to* "$>$", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d$, $p \neq k$. *If* $a_{2k} > 0$, $a_{2k} \geqslant b_{2k} \geqslant 0$, $a_{2k} \geqslant c_{2k} \geqslant 0$, *and the value for* a_{2k} *is less than or equal to the maximum length of the right-hand-side intervals of the interval equations in the variable interval equation*, *then the variable interval equation is* (Rel. 3.4)-*integer solvable if and only if the variable interval equation*

$$a_1X_1 + \cdots + a_{2k-1}X_{2k-1} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n$$
$$= \left[b_0 + \sum_{r=1}^{n} b_r x_r + (b_{2k} - a_{2k})(X_{2k-1} - 1),\right.$$
$$\left. c_0 + \sum_{r=1}^{n} c_r x_r + (c_{2k} - a_{2k})\left(P_{2k,0} + \sum_{s=1}^{2k-1} P_{2k,s}X_s\right)\right] \quad \text{for } r \neq 2k \tag{3.7}$$

is (Rel. 3.5)-*integer solvable.*

Proof. Refer to Theorems 3.1 and 3.2. □

Theorem 3.5. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* "$>$", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d$, $p \neq k$. *If* $a_{2k} < 0$, $a_{2k} \leqslant b_{2k} \leqslant 0$, $a_{2k} \leqslant c_{2k} \leqslant 0$, *and the negative value for* a_{2k} *is less than or equal to the length of the right-hand-side interval on one of the interval equations in the variable interval equation, then the variable interval equation is* (Rel. 3.2)-*integer solvable if and only if the variable interval equation*

$$a_1X_1 + \cdots + a_{2k-1}X_{2k-1} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n$$
$$= \left[b_0 + \sum_{r=1}^{n} b_r x_r + (b_{2k} - a_{2k})\left(P_{2k,0} + \sum_{s=1}^{2k-1} P_{2k,s}X_s\right),\right.$$
$$\left. c_0 + \sum_{r=1}^{n} c_r x_r + (c_{2k} - a_{2k})(X_{2k-1} - 1)\right] \quad \text{for } r \neq 2k \tag{3.8}$$

is (Rel. 3.5)-*integer solvable.*

Proof. Refer to Theorem 3.2. □

Lemma 3.5. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* "$>$", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d$, $p \neq k$. *If* $a_{2k} < 0$, $a_{2k} \leqslant b_{2k} \leqslant 0$, $a_{2k} \leqslant c_{2k} \leqslant 0$, *and the negative value for* a_{2k} *is less than or equal to the maximum length of the right-hand-side intervals of the interval equations in the variable interval equation, then the variable interval equation is* (Rel. 3.2)-*integer solvable if and only if the variable interval equation*

$$a_1X_1 + \cdots + a_{2k-1}X_{2k-1} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n$$
$$= \left[b_0 + \sum_{r=1}^{n} b_r x_r + (b_{2k} - a_{2k})\left(P_{2k,0} + \sum_{s=1}^{2k-1} P_{2k,s}X_s\right),\right.$$
$$\left. c_0 + \sum_{r=1}^{n} c_r x_r + (c_{2k} - a_{2k})(X_{2k-1} - 1)\right] \quad \text{for } r \neq 2k \tag{3.9}$$

is (Rel. 3.5)-*integer solvable.*

Proof. Refer to Theorems 3.1 and 3.2. □

Let us use the following example to explain the power of the GDVI test based on Lemmas 3.4 and 3.5, when it is applied to deal with a data dependence problem for a linear equation with *variable* bounds under a given direction vector.

Consider the equation

$$X_1 - X_2 - X_3 + X_4 = 0$$

subject to the bounds

$$1 \leqslant X_1 \leqslant 100,\ 1 \leqslant X_2 \leqslant 100,\ 1 + X_1 \leqslant X_3 \leqslant 100 + X_1 \text{ and } 1 + X_2 \leqslant X_4 \leqslant 100 + X_2,$$

and the limits of a direction vector

$$X_1 > X_2 \quad \text{and} \quad X_3 > X_4.$$

According to the constraints of (2.2) denoted in Definition 2.2, the constraints for X_1, X_2, X_3 and X_4 will be redefined by $1 \leqslant X_1 \leqslant 100$, $1 \leqslant X_2 \leqslant X_1 - 1$, $1 + X_1 \leqslant X_3 \leqslant 100 + X_1$, and $1 + X_2 \leqslant X_4 \leqslant X_3 - 1$. If the GDVI test is used to resolve the problem, then in light of Definitions 3.2a–3.2c the set of all the interval equations Ψ is equal to

$$\{X_1 - X_2 - X_3 + X_4 = [0, 0] \mid 1 \leqslant X_1 \leqslant 100,\ 1 \leqslant X_2 \leqslant X_1 - 1,\\ 1 + X_1 \leqslant X_3 \leqslant 100 + X_1 \text{ and } 1 + X_2 \leqslant X_4 \leqslant X_3 - 1\}.$$

The set Ω for the length of the right-hand-side interval on every interval equation in the set Ψ is $\{1\}$. Therefore, the maximum element in the set Ω is one. It is obvious from Definitions 3.2a–3.2c that the set Ψ is integer solvable if the only interval equation in the set Ψ is integer solvable. The coefficient for X_4 satisfies the assumption of Lemma 3.4: (1) $1 > 0$, (2) $1 \geqslant 0 \geqslant 0$, (3) $1 \geqslant 0 \geqslant 0$ and (4) the value of the coefficient is equal to one. Lemma 3.4 is applied towards moving the term X_4 to the right-hand side of the only interval equation in the set Ψ. The variable interval equation (i.e., the new set Ψ_1) in light of Lemma 3.4 and Definitions 3.2a–3.2c is

$$\{X_1 - X_2 - X_3 = [1 - X_3, -1 - X_2] \mid 1 \leqslant X_1 \leqslant 100,\ 1 \leqslant X_2 \leqslant X_1 - 1 \text{ and}\\ 1 + X_1 \leqslant X_3 \leqslant 100 + X_1\}.$$

Now the set Ω_1 for the length of the right-hand-side interval on every interval equation in the set Ψ_1 is equal to $\{X_3 - X_2 - 1 \mid 1 \leqslant X_1 \leqslant 100,\ 1 \leqslant X_2 \leqslant X_1 - 1$ and $1 + X_1 \leqslant X_3 \leqslant 100 + X_1\}$. The maximum element computed by the Banerjee algorithm in the set Ω_1 is 198. When the maximum element is equal to 198, the values for X_1, X_2 and X_3 are equal to 100, 1 and 200, respectively. Because $X_1 = 100$, $X_2 = 1$ and $X_3 = 200$, so $-101 = [-199, -2]$ $(-199 \leqslant -101 \leqslant -2)$ hold. Therefore, there exists a constant interval equation in the set Ψ_1 satisfying the given limitations. The coefficient for X_3 satisfies the assumption of Lemma 3.5: (1) $-1 < 0$, (2) $-1 \leqslant -1 \leqslant 0$, (3) $-1 \leqslant 0 \leqslant 0$ and (4) the negative value of the coefficient is less than 198. Lemma 3.5

is applied towards moving the term X_3 to the right-hand side of the interval equation in the set Ψ_1. The variable interval equation (i.e., the new set Ψ_2) in light of Lemma 3.5 and Definitions 3.2a–3.2c is

$$\{X_1 - X_2 = [1, 99 + X_1 - X_2] \mid 1 \leqslant X_1 \leqslant 100 \text{ and } 1 \leqslant X_2 \leqslant X_1 - 1\}.$$

Now the set Ω_2 for the length of the right-hand-side interval on every interval equation in the set Ψ_2 is equal to $\{99 + X_1 - X_2 \mid 1 \leqslant X_1 \leqslant 100 \text{ and } 1 \leqslant X_2 \leqslant X_1 - 1\}$. The maximum element computed by the Banerjee algorithm in the set Ω_2 is 198. When the maximum element is equal to 198, the values for X_1 and X_2 are equal to 100 and 1, respectively. Because $X_1 = 100$ and $X_2 = 1$, so $99 = [1, 198]$ $(1 \leqslant 99 \leqslant 198)$ hold. Therefore, there exists a constant interval equation in the set Ψ_2 satisfying the given limitations. The coefficient for X_2 satisfies the assumption of Lemma 3.5: (1) $1 < 0$, (2) $-1 \leqslant 0 \leqslant 0$, (3) $-1 \leqslant -1 \leqslant 0$ and (4) the negative value of the coefficient is less than 198. Lemma 3.5 is applied towards moving the term X_2 to the right-hand side of the interval equation in the set Ψ_2. The variable interval equation (i.e., the new set Ψ_3) in light of Lemma 3.5 and Definitions 3.2a–3.2c is

$$\{X_1 = [2, 99 + X_1] \mid 1 \leqslant X_1 \leqslant 100\}.$$

Now the set Ω_3 for the length of the right-hand-side interval on every interval equation in the set Ψ_3 is equal to $\{98 + X_1 \mid 1 \leqslant X_1 \leqslant 100\}$. The maximum element computed by the Banerjee algorithm in the set Ω_3 is 198. When the maximum element is equal to 198, the value for X_1 is equal to 100. Because $X_1 = 100$, so $100 = [2, 199]$ $(2 \leqslant 100 \leqslant 199)$ hold. Therefore, there exists a constant interval equation in the set Ψ_3 satisfying the given limitations. The coefficient for X_1 satisfies the assumption of Lemma 3.4: (1) $1 > 0$, (2) $1 \geqslant 0 \geqslant 0$, (3) $1 \geqslant 1 \geqslant 0$ and (4) the value of the coefficient is less than 198. Lemma 3.4 is again used to move the term X_1 to the right. The new set Ψ_4 is

$$\{0 = [-98, 99]\}.$$

According to Definitions 3.2a–3.2c, the set Ψ_4 is equal to $\{0 = [-98, 99]\}$. The expression of the left-hand side on the only interval equation in the set Ψ_4 is reduced to zero items. The only interval equation in the set Ψ_4 is integer solvable because $-98 \leqslant 0 \leqslant 99$ is true. Therefore, the GDVI test *concludes* that there are integer solutions.

3.4. Transformation of variable interval equation with "=" in direction vector

Two variables in (3.1) will be merged into a single variable if they refer to the same index variable and are related in dependence direction by a direction vector "=". The bounds for the single variable will be unchanged. Theorems 3.6 and 3.7 and Lemmas 3.6 and 3.7 are derived to describe the way of variable-interval-equation to variable-interval-equation transformations under variable limits as well as a specific direction vector $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, where d is the number of common loops, θ_k is equal to "=", $1 \leqslant k \leqslant d$, and θ_p are equal to "$*$" for $1 \leqslant p \leqslant d$, $p \neq k$.

Theorem 3.6. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* "=", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "*" *for* $1 \leqslant p \leqslant d, p \neq k$. *If* $(a_{2k-1} + a_{2k}) > 0, (a_{2k-1} + a_{2k}) \geqslant (b_{2k-1} + b_{2k}) \geqslant 0, (a_{2k-1} + a_{2k}) \geqslant (c_{2k-1} + c_{2k}) \geqslant 0$, *and the value for* $(a_{2k-1} + a_{2k})$ *is less than or equal to the length of the right-hand-side interval on one of the interval equations in the variable interval equation*, *then the variable interval equation is* (Rel. 3.6)-*integer solvable*, *where* (Rel. 3.6) *is*

$$\left(\begin{array}{l} P_{2k,0} + \sum_{s=1}^{2k-1} P_{2k,s} X_s \leqslant X_{2k-1} = X_{2k} \leqslant Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s} X_s, \\ \quad \text{where } 1 \leqslant k \leqslant d;\ P' \leqslant X_{2p-1} \leqslant Q' \text{ and } P'' \leqslant X_{2p} \leqslant Q'' \text{ for } 1 \leqslant p \leqslant d,\ p \neq k, \\ \quad \text{the values of } P', Q', P'', Q'' \text{ will depend on the dependence direction } \theta p; \\ \quad P_{r,0} + \sum_{s=1}^{r-1} P_{r,s} X_s \leqslant X_r \leqslant Q_{r,0} + \sum_{s=1}^{r-1} Q_{r,s} X_s \text{ for } 2d+1 \leqslant r \leqslant n \end{array} \right), \tag{Rel.3.6}$$

if and only if the variable interval equation

$$\begin{aligned} & a_1 X_1 + \cdots + a_{2k-2} X_{2k-2} + a_{2k+1} X_{2k+1} + \cdots + a_n X_n \\ & \quad = \left[b_0 + \sum_{r=1}^{n} b_r x_r + (b_{2k-1} + b_{2k} - a_{2k-1} - a_{2k}) \left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s} X_s \right), \right. \\ & \qquad \left. c_0 + \sum_{r=1}^{n} c_r x_r + (c_{2k-1} + c_{2k} - a_{2k-1} - a_{2k}) \left(P_{2k,0} + \sum_{s=1}^{2k-1} P_{2k,s} X_s \right) \right] \\ & \qquad \text{for } r \neq 2k-1 \text{ and } 2k \end{aligned} \tag{3.10}$$

is (*Rel.* 3.7)-*integer solvable*, *where* (*Rel.* 3.7) *is*

$$\left(\begin{array}{l} P' \leqslant X_{2p-1} \leqslant Q' \text{ and } P'' \leqslant X_{2p} \leqslant Q'' \text{ for } 1 \leqslant p \leqslant d,\ p \neq k, \\ \quad \text{the values of } P', Q', P'', Q'' \text{ will depend on the dependence direction } \theta_p; \\ \quad P_{r,0} + \sum_{s=1}^{r-1} P_{r,s} X_s \leqslant X_r \leqslant Q_{r,0} + \sum_{s=1}^{r-1} Q_{r,s} X_s \text{ for } 2d+1 \leqslant r \leqslant n \end{array} \right). \tag{Rel.3.7}$$

Proof. Refer to Theorem 3.2. □

Lemma 3.6. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* "=", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "*" *for* $1 \leqslant p \leqslant d, p \neq k$. *If* $(a_{2k-1} + a_{2k}) > 0,\ (a_{2k-1} + a_{2k}) \geqslant (b_{2k-1} + b_{2k}) \geqslant 0,\ (a_{2k-1} + a_{2k}) \geqslant (c_{2k-1} + c_{2k}) \geqslant 0$, *and the negative value for* $(a_{2k-1} + a_{2k})$ *is less than or equal to the maximum length of* *the right-hand-side interval on one of the interval equations in the variable interval equation, then the variable interval equation is* (Rel. 3.6)-*integer solvable if and only if the variable interval equation* (3.10) *is* (Rel. 3.7)-*integer solvable.*

Proof. Refer to Theorem 3.2. □

Theorem 3.7. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* "$=$", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d, p \neq k$. *If* $(a_{2k-1} + a_{2k}) < 0$, $(a_{2k-1} + a_{2k}) \leqslant (b_{2k-1} + b_{2k}) \leqslant 0$, $(a_{2k-1} + a_{2k}) \leqslant (c_{2k-1} + c_{2k}) \leqslant 0$, *and the negative value for* $(a_{2k-1} + a_{2k})$ *is less than or equal to the length of the right-hand-side interval on one of the interval equations in the variable interval equation, then the variable interval equation is* (Rel. 3.6)-*integer solvable if and only if the variable interval equation*

$$\begin{aligned}
&a_1X_1 + \cdots + a_{2k-2}X_{2k-2} + a_{2k+1}X_{2k+1} + \cdots + a_nX_n \\
&= \left[b_0 + \sum_{r=1}^{n} b_r x_r + (b_{2k-1} + b_{2k} - a_{2k-1} - a_{2k}) \left(P_{2k,0} + \sum_{s=1}^{2k-1} P_{2k,s} X_s \right), \right. \\
&\qquad \left. c_0 + \sum_{r=1}^{n} c_r x_r + (c_{2k-1} + c_{2k} - a_{2k-1} - a_{2k}) \left(Q_{2k,0} + \sum_{s=1}^{2k-1} Q_{2k,s} X_s \right) \right] \\
&\qquad \textit{for } r \neq 2k-1 \textit{ and } 2k
\end{aligned} \tag{3.11}$$

is (Rel. 3.7)-*integer solvable.*

Proof. Refer to Theorem 3.2. □

Lemma 3.7. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops,* θ_k *is equal to* "$=$", $1 \leqslant k \leqslant d$, *and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d$, $p \neq k$. *If* $(a_{2k-1} + a_{2k}) < 0$, $(a_{2k-1} + a_{2k}) \leqslant (b_{2k-1} + b_{2k}) \leqslant 0$, $(a_{2k-1} + a_{2k}) \leqslant (c_{2k-1} + c_{2k}) \leqslant 0$, *and the negative value for* $(a_{2k-1} + a_{2k})$ *is less than or equal to the maximum length of the right-hand-side interval on one of the interval equations in the variable interval equation, then the variable interval equation is* (Rel. 3.6)-*integer solvable if and only if the variable interval equation* (3.11) *is* (Rel. 3.7)-*integer solvable.*

Proof. Refer to Theorem 3.2. □

Two variables referring to the same loop index variable in the variable interval equation (3.1) will be merged into a single variable if they are related by a direction vector "$=$". The bounds for the single variable will be unchanged. Theorems 3.6 and 3.7 and Lemmas 3.6 and 3.7 can be used to determine integer solutions for the variable interval equation (3.1) with the bounds of (1.2) and specific direction vectors containing "$=$" if the coefficient $(a_{2k-1} + a_{2k})$ for the single variable in the equation is small enough to justify the movement of the item to the right. If $(a_{2k-1} + a_{2k}) > 0$,

$(a_{2k-1}+a_{2k}) \geqslant (b_{2k-1}+b_{2k}) \geqslant 0$, $(a_{2k-1}+a_{2k}) \geqslant (c_{2k-1}+c_{2k}) \geqslant 0$, and the value for $(a_{2k-1}+a_{2k})$ is less than or equal to one of the elements in the set Ω in light of the assumption of Theorem 3.6 and Lemma 3.6, then the value is actually equivalent to the small enough value to justify the movement of the item to the right. If $(a_{2k-1}+a_{2k}) < 0$, $(a_{2k-1}+a_{2k}) \leqslant (b_{2k-1}+b_{2k}) \leqslant 0$, $(a_{2k-1}+a_{2k}) \leqslant (c_{2k-1}+c_{2k}) \leqslant 0$, and the negative value for $(a_{2k-1}+a_{2k})$ is less than or equal to one of the elements in the set Ω in light of the assumption of Theorem 3.7 and Lemma 3.7, then the value is actually equivalent to the small enough value to justify the movement of the item to the right. On the other hand, Theorems 3.6 and 3.7 and Lemmas 3.6 and 3.7 are inapplicable towards ascertaining integer solutions of the variable interval equation (3.1) if the absolute values of the coefficients for all the items in the variable interval equation (3.1) are greater than the maximum element in the set Ω. The Banerjee algorithm can be employed to determine the maximum element in the set Ω.

The GDVI test generates three possible results. The first result of 'yes' means that Eq. (3.1) has integer solutions, and the second result of 'no' means that there are no integer solutions. The third value of 'maybe', on the other hand, shows that the equation has a solution which satisfies the limits on all the variables which has been moved to the right-hand side of the equation, and might still have a solution which satisfies the limits on the rest of the variables.

The GDVI test produces a result of 'maybe' because there are no longer any coefficients with small enough values to justify their movement to the right. In the case, it is prudent to complete the "step-by-step *Banerjee algorithm*" anyway, i.e., to complete the computation of the Banerjee bounds, L_b and U_b, and to test for $[L_b, U_b] \cap [L, U] = \emptyset$, where $[L, U]$ is the right-hand side of the variable interval equation (3.1) after Theorem 3.1 has been applied as many times as possible. This is to imply that the GDVI test is always at least as efficient and accurate as the Banerjee algorithm.

The GDVI test can be viewed as involving the term-by-term computation of the Banerjee bounds. That is, the Banerjee-bound-computation component of the GDVI test costs at most the cost of a *single* Banerjee algorithm. If the GDVI test arrives at a definitive result (for example, all of the integers in between the intervals of the interval equations in the set Ψ are not divisible by the greatest common divisor of the left-hand-side coefficients of (3.1)) before all terms have been moved to the right-hand side of the variable interval equation, then the Banerjee-bound-computation component of the GDVI test costs even less.

3.5. Variable interval-equation transformation using the GCD test

It is obvious from Lemmas 3.2–3.7 that one variable in Eq. (3.1) can be moved to the right if the coefficient of the variable has a small enough value to justify the movement of the variable to the right. If all coefficients for variables in Eq. (3.1) have no sufficiently small values to justify the movements of variables to the right, then Lemmas 3.2–3.7 cannot be applied. In the following, Lemma 3.8, extended from Theorem 2.2, describes the new transformation using the GCD test which frequently enables one or more additional variables to be moved.

Lemma 3.8. *Given a variable interval equation* (3.1) *subject to the constraints of* (1.2) *and a specific direction vector* $\vec{\theta} = (\theta_1, \ldots, \theta_d)$, *where d is the number of common loops, and* θ_p *are equal to* "$*$" *for* $1 \leqslant p \leqslant d$. *Let* $g = \gcd(a_1, \ldots, a_n)$. *The variable interval equation* (3.1) *is* $(P_1, Q_1; \ldots; P_n, Q_n)$-*integer solvable if and only if the variable interval equation*

$$(a_1/g)X_1 + \cdots + (a_{2d}/g)X_{2d} + \cdots + (a_n/g)X_n = [\lceil L/g \rceil, \lfloor U/g \rfloor] \qquad (3.12)$$

is $(P_1, Q_1; \ldots; P_n, Q_n)$-*integer solvable, where* P_k, Q_k, $1 \leqslant k \leqslant n$, *in* $(P_1, Q_1; \ldots; P_n, Q_n)$ *refer to the bounds of* X_k *but may be redefined due to dependence direction.*

Proof. Refer to [11]. □

Lemma 3.8 guarantees to always perform at least as well as (and sometimes better than) a combination of the GCD test and the Banerjee algorithm at no more than their cost (and sometimes at a lower cost). In the worst case, the GDVI test consists of n GCD tests, where n is the number of variables in Eq. (3.1). In actual practice, it requires frequently no more than one.

3.6. Time complexity

The main phases to the GDVI test are:

1. finding a small enough coefficient to justify the movement of a term to the right-hand side;
2. changing the expression of the right-hand side on the equation due to the movement of a term to the right;
3. using the GCD test to reduce coefficients of each variable in Eq. (3.1).

A small enough coefficient is easily found according to Theorems 3.1–3.7, and Lemmas 3.2–3.7. It is obvious that the worst-case time complexity to searching such a coefficient is $O(n^2 + y * n)$ (n^2 accounts for finding the maximum length of variable interval and $y * n$ is due to search of proper term) in light of Theorem 3.1 and Lemmas 3.2–3.7, where n is the number of variables in the equation and y is a constant. The number of looking for all small enough coefficients in the equation is at most n times because the number of terms moved in the equation is at most n terms. Thus, the worst-case time complexity to finding all small enough coefficients in the variable interval equation is concluded to be $O(n^3 + n^2 * y)$.

The expression of the right-hand side on the variable interval equation is changed according to Lemmas 3.2–3.7 because an item on the left-hand side of the equation is moved to the right. The cost of changing the expression of the right-hand side on the equation according to Theorem 3.1 and Lemmas 3.2–3.7 is actually equivalent to a *single* term computation of the *Banerjee algorithm*. The worst-case time complexity to a single term computation of the Banerjee algorithm is $O(n)$. Thus, the worst-case time complexity to modifying the expression of the right-hand side on a variable interval equation is deduced to be $O(n)$. The number to modifying the expression of the right-hand side on the equation is at most n times because the number of terms is at most n. Therefore, the worst-case time complexity for changing all the expressions of the right-hand side on the variable interval equation is $O(n^2)$. If all coefficients in Eq. (3.1) have no absolute values of 1, then Lemma 3.8 employs the GCD test to reduce all coefficients to obtain a small enough coefficient. In the worst cases, the GDVI test contains n GCD tests. Ref. [2] shows that a large percentage of all coefficients have absolute values of 1 in one-dimensional array references with linear subscripts in real programs. Therefore, the GCD test is very seldom to be applied to reduce the coefficients in the equations of real programs, implying the time complexity for the GCD test can be ignored. The worst-case time complexity to the GDVI test is thus derived to be $O(n^3 + y * n^2 + n^2)$. The worst-case time complexity of the original DVI test is $O(n^2 * y + n * y)$ [1]. That is, the GDVI test has slightly decreased efficiency than that of the original DVI test because the number of variables, n, in the equation tested is generally very small.

4. Experimental results

We tested the effect of the GDVI test and compared it with other tests through performing experiments on Personal Computer Intel 80486 to the codes cited from five numerical packages EISPACK, LINPACK, Parallel loops, Livermore loops and Vector loops [8,9,12–14]. In our experiments, if the stride of the loop is a positive constant, then whenever a loop lower limit is not known 1 is assumed, and whenever the loop upper limit is not known 100 is assumed. The choices of 1 and 100 as the loop lower and upper bounds are arbitrary. The stride of the loop was defined to be 1 if it was not an integer constant. We did not perform induction variable substitution in the experiments. The codes tested include about 37 000 lines of statements involving 205 subroutines and 12 152 pairs of one-dimensional array references consisting of the same pair of array references but with different direction vectors. The results are shown in Table 1. The DVI test detected that there were *definite* (*yes* or *no*) results for 6629 pairs of one-dimensional array references with constant bounds under any given direction vectors. The GDVI test is only applied to those one-dimensional arrays with linear subscripts under variable bounds as well as any given direction vectors. It found that there were 2124 pairs with *definite* results. Therefore, there were 8753 *definite* results obtained based on interval testing approach.

The improvement rate of the GDVI test can be affected by two factors. First, the frequency of one-dimensional array references with linear subscripts subject to

Table 1
Comparisons of testing capability among the DVI test, the GDVI test, the Power test, and the Omega test for 12 152 pairs of benchmark statements

Dependence testing	Constant bounds		Variable bounds		Overall		Success rate (%)
	Definitive	Maybe	Definitive	Maybe	Definitive	Maybe	
DVI test	6629	2821			6629	2821	54.5
GDVI test			2124	578	2124	578	17.5
Power test	4292	5158	1416	1286	5708	6444	47.0
Omega test	6629	2821	2124	578	8753	3399	72.0

constant and variable limits. Second, the "success rate" of the GDVI test, by which we mean how often the GDVI test detects a case where there is a definite result. Let b be the number of the one-dimensional arrays found in our experiments, and let c be the number that is detected to have definite results from the one-dimensional arrays with linear subscripts subject to constant and variable limits and any given direction vectors. Thus the success rate is denoted to be equal to c/b. In our experiments, 12 152 pairs of one-dimensional array references were found, and 2124 of them were found to have definite results to variable bounds. So the improvement rate of the GDVI test over the DVI test in our experiments was about 17.5%.

The Power test and Omega test were also tested to resolve those 12 152 pairs of one-dimensional array references. The Power test concluded that there were 1416 pairs with *definite* results for variable limits and 4292 pairs with *definite* results to constant constraints. This indicates that the Power test is not as accurate as the interval testing approach. Whereas, the Omega test, as shown in Table 1, was found to give the same accurate results as the interval testing approach when it was used to handle dependence testing of one-dimensional array references with constant and variable bounds under any given direction vectors.

The execution efficiency for these test approaches was also compared. Let K_{D}, K_{GD}, K_{P} and K_{O} be the execution time to treat data dependence problem of a one-dimensional array for the DVI test, the GDVI test, the Power test, and the Omega test, subsequently. Table 2 shows the computing speedups the DVI test and the GDVI test over the Power test and the Omega test for those 12 152 pairs of array references. It is very clear that the DVI test and the GDVI test are much superior to the Power test and the Omega test in terms of analyzing efficiency.

The superiority of testing efficiency of the GDVI test over that of the Omega test for the stated dependence problem can also be deduced from time complexity analysis. The Omega test based on the least remainder algorithm, a variation of Euclid's algorithm, and Fourier's elimination method [7,10] consists of three major computations: eliminating equality constraints, eliminating variables in inequality constraints, and finding integer solutions (that is an integer programming problem). The time complexities for these steps are $O(mn \log|c| + mnp + mn)$, $O(n^2s^2)$ and $O(k^n)$ [7,10,16], respectively, where m, n, c, p, s, k denote the number of equality constraints,

Table 2
The computing speedups of the DVI test and the GDVI test over the Power test and the Omega test for 12 152 pairs of benchmark statements

	Bounds of tested loops	Speed-up	Total number of subroutines involved
$K_{\mathrm{P}}/K_{\mathrm{D}}$	Constant	4.5–5.0	40
		9.5–18.0	126
$K_{\mathrm{P}}/K_{\mathrm{GD}}$	Variable	6.5–11.0	6
		13.0–22.0	33
$K_{\mathrm{O}}/K_{\mathrm{D}}$	Constant	4.5–9.5	156
		10.0–17.5	10
$K_{\mathrm{O}}/K_{\mathrm{GD}}$	Variable	3.0–12.0	36
		18.0–21.0	3

the number of variables, the coefficient with the largest absolute value in equality constraints, the number of passes to eliminate all the variables that become unbound, the number of inequality constraints, and the absolute value of coefficient of variable in inequality constraints, subsequently. So the overall time complexity of the Omega test is $O(mn \log |c| + mnp + mn + n^2s^2 + k^n)$. Obviously, the GDVI test is significantly superior to that of the Omega test in terms of testing efficiency. In [7] it is reported that the Omega test has *exponential* worst-case time complexity. Wolf [6,7] and Triolet [5] also found that Fourier–Motzkin variable elimination for dependence testing takes from 22 to 28 times longer than Banerjee method, a part of the GDVI test.

5. Conclusions

The GDVI test proposed in this paper extends the dependence testing range of one-dimensional array references to linear subscripts with variable bounds under any given direction vectors, enhancing significantly data dependence analysis capability of the DVI test. The GDVI test defines some conditions under which dependence equations of linear subscripts with variable bounds under any given direction vectors can be continuously tested if integer solutions exist. In short, the GDVI test is exactly equivalent to a version of the DVI test that combines the Banerjee algorithm and the GCD test. The Banerjee algorithm is employed to deduce maximum bound to a linear expression of which the variables are with variable constraints and any given direction vectors [3].

The Power test is a combination of Fourier–Motzkin variable elimination with an extension of Euclid's GCD algorithm [4,6]. The Omega test combines new methods for eliminating equality constraints with an extension of Fourier–Motzkin variable elimination to integer programming [7]. These two tests have currently the highest precision and the widest applicable range in the field of data dependence testing for array references with linear subscripts. However, the cost of the two tests is very expensive because the worst-case of Fourier–Motzkin variable elimination is exponential in the number of free variables [4,6,7]. Banerjee [10] indicated that the Omega test is a precise but inefficient method. According to our experiments, the efficiency and the precision of the interval testing approach are much better than those of the Power test. Whereas, the interval testing approach shares the same accuracy with the Omega test but outperforms significantly the Omega test in testing efficiency.

The GDVI test extends the applicable range of the DVI test and, according to the time complexity analysis, only slightly lowers the efficiency of the DVI test. Therefore, the GDVI test seems to be a practical scheme to analyze data dependence for one-dimensional arrays with linear references.

References

[1] X. Kong, D. Klappholz, K. Psarris, The direction vector i test, IEEE Transaction on Parallel and Distributed Systems 4 (11) (1993) 1280–1290.

[2] Z. Shen, Z. Li, P.-C. Yew, An empirical study of Fortran programs for parallelizing compilers, IEEE Transaction on Parallel and Distributed Systems 1 (3) (1992) 356–364.
[3] U. Banerjee, in: Dependence Analysis for Supercomputing, Kluwer Academic Publishers, Dordrecht, 1988.
[4] M. Wolfe, C.-W. Tseng, The power test for data dependence, IEEE Transaction on Parallel and Distributed Systems 3 (5) (1992) 591–601.
[5] R. Triolet, F. Irigoin, P. Feautrier, Direct parallelization of call statements, in: Proc. SIGPLAN Symp. Compiler Constructiton, Palo Alto, CA, 1986, pp. 176–185.
[6] M. Wolfe, in: High Performance Compiler for Parallel Computing, Addison-Wesley, Reading, MA, 1996.
[7] W. Pugh, A practical algorithm for exact array dependence analysis, Communication of the ACM 35 (8) (1992) 102–114.
[8] B.J. Smith, in: Matrix Eigensystem Routines – Eispack Guide, Springer, Heidelberg, 1976.
[9] W. Blume, R. Eigenmann, Performance analysis of parallelizing compilers on the perfect benchmark® programs, IEEE Transaction on Parallel and Distributed Systems 3 (6) (1992) 643–656.
[10] U. Banerjee, in: Dependence Analysis, Kluwer Academic Publishers, Norwell, Massachusetts, 1997.
[11] W.-L. Chang, C.-P. Chu, The extension of the i test, Parallel Computing 24 (14) (1998) 2101–2127.
[12] J. Dongarra, M. Furtney, S. Reinhardt, J. Russell, Parallel loops – a test suite for parallelizing compilers: description and example results, Parallel Computing 17 (1991) 1247–1255.
[13] J.M. Levesque, J.W. Williamson, in: A Guidebook to Fortran on Supercomputing, Academic Press, New York, 1989.
[14] D. Levine, D. Callahan, J. Dongarra, A comparative study of automatic vectorizing compilers, Parallel Computing 17 (1991) 1223–1244.
[15] X. Kong, D. Klappholz, K. Psarris, The i test, IEEE Transaction on Parallel and Distributed Systems 2 (3) (1991) 342–349.
[16] U. Banerjee, in: Loop Transformations for Restructuring Compilers: The Foundations, Kluwer Academic Publishers, Dordrecht, 1993.