

OPERATION SEQUENCE BASED SIMILARITY COEFFICIENT FOR RMS

GOYAL, K[apil] K[umar]; JAIN, P[ramod] K[umar] & JAIN, M[adhu]

Abstract: Reconfigurable manufacturing system is characterized by rapid and cost effective adaptation to the unscheduled and un-predictable changes in the market requirements through changeable functionality and scalable capacity around a part family. The design of RMS layout for catering to the product family requires an effective approach to group the parts based on operation sequence. In past the developed similarity coefficients suffer from the drawbacks. In the present work authors propose a BMIM (by passing moves and idle machines) similarity coefficient based on longest operation sequence. The superiority of the developed similarity coefficient is demonstrated through an illustration.

Keywords: reconfigurable manufacturing system, reconfigurability, reconfigurable machine tool, operation sequence, similarity coefficient, part family

1. INTRODUCTION

From the beginning of 21st Century, manufacturing has entered into a new era in which all the manufacturing enterprises must face the customer driven markets and the fierce competition. To stay competitive the manufacturers must use systems that not only produce goods with high productivity, but also allow for rapid response to market changes and customer needs. The reconfigurable manufacturing system (RMS) is envisaged to cater the volatile markets and the required rapid reconfigurations by their modular structure [1].

The RMS is built around a part family i.e. offering the exact capacity and functionality needed to process a part family and behaves as DMS during the production phase and can readily be reconfigured according to the new manufacturing requirements [2]. Therefore like DMS the most appropriate layout for the RMS is also the flow line layout to support mass manufacturing at the competitive cost. In most of the RMS modelling approaches flow line layout has been adopted [3-6]. Further in designing the flow line layout, operation sequence similarity plays the pivotal role to reduce the overall cost and reconfigurations required while producing set of part families, which in turn can be only achieved through reducing the machine idleness along with the material handling effort. In the previous works on part family formation in RMS, operation sequence similarity has not been considered to the best knowledge of the authors [7-9]. The approaches developed in the past based on operation sequence similarity also suffer from various limitations and shortcomings which has been described later in the section 5.

In nutshell, various highly referred best similarity/dissimilarity coefficients suffer from the limitations in one aspect or the other. Therefore authors

have developed a novel LCS and shortest composite supersequence (SCS) based BMIM (Bypassing Moves and Idle Machines) similarity coefficient which considers the bypassing moves and idleness of machines accurately.

2. FINDING LCS AND SCS

In the first phase of computing the similarity coefficient between two parts, the maximum number of identical operations from both the operation sequences which can be performed in order of precedence relationship are to be listed. This list of common operations of both the operation sequences following precedence constraints is formally known as longest common subsequence. For example, consider two operation sequences $A = \{a d c b\}$ and $B = \{f d e c h b\}$. The $\{d b\}$ is a subsequence of both the sequences A and B but it is not the longest common subsequence. The LCS of A and B is $\{d c b\}$.

The edit distance in the LCS problem is defined by the minimum number of elementary transformation operations (insertions, deletions and substitutions) required to transform the source string into target string [10]. The classic dynamic programming solution to LCS problem was invented by [11]. Since then the longest common subsequence of two strings has been studied extensively by various researchers [12-14]. For a comprehensive comparison of the well-known algorithms for LCS problem the contributions of [15] may be referred. The pseudo code based on [11], which applies the classical dynamic programming to find the LCS between two operation sequences, is presented below.

function $LCS_length(A, B)$

```

1  m ← |A|; n ← |B|;
2  for i ← 1 to m
3      C[i, 0] ← 0; end
4  for j ← 0 to n
5      C[0, j] ← 0; end
6  for i ← 1 to m
7      for j ← 1 to n
8          if A[i] = B[j]
9              C[i, j] ← C[i - 1, j - 1] + 1;
10             else if C[i - 1, j] ≥ C[i, j - 1]
11                 C[i, j] ← C[i - 1, j];
12             else
13                 C[i, j] ← C[i, j - 1];
14             end
15

```

```

16 end
17 end
18 return C[ m, n ];

```

function *LCS_print*(A, B, C)

```

1  i ← |A|; j ← |B|;
2  while ( i ≠ 0 & j ≠ 0) do
3    if A[ i ]=B[ j ]
4      print A[ i ];
5      i ← i - 1; j ← j - 1;
6    elseif C[ i, j-1 ] > C[ i-1, j ]
7      j ← j-1;
8    else
9      i ← i-1;
10 end
11 end

```

The pseudo code of the function *LCS_length* computes the length of LCS between two operation sequences A[1..m] and B[1..n] by computing the lengths of LCS between A[1..i] and B[1..j] $\forall 1 \leq i \leq m$ and $1 \leq j \leq n$, and stores it in C[i j]. C[m n] will contain the length of the LCS of operation sequences A and B. Further in the function *LCS_print*, the LCS is extracted from the matrix C.

The shortest composite supersequence (SCS) is an arrangement obtained from LCS of both sequences, which is the shortest possible length of a sequence which accommodates all the operations of both the sequences following their precedence constraints. The length of SCS between two operation sequences *x* and *y* may be obtained after finding the LCS using:

$$|SCS_{xy}| = |x| + |y| - |LCS_{xy}| \quad (1)$$

3. MATHEMATICAL MODEL FOR DETERMINING BMIM SIMILARITY

The order of appending the left out operations from both the operation sequences into the LCS for obtaining the SCS may affect the material flow. Therefore the left out operations of an operation sequence to be appended in the LCS to obtain SCS are divided into two categories. First category is the set of operations which are to be appended in between the LCS. In the second category, the operations are the set of operations from each operation sequence to be appended before or after the LCS. The first category does not affect the number of bypassing moves, but the position of appending second category operations from both the operation sequences to form SCS may affect the number of bypassing moves. Here an SCS offering minimum number of bypassing moves is to be determined through the proposed algorithm. In the second category, there are two sets of operations from each operation sequence, one set to be appended before the LCS and another set to be appended after the LCS. The following mathematical formulation is adopted to find the minimum number of bypassing moves and idle machines to compute the BMIM similarity coefficient. Following are the notations for mathematical modelling of the present work:

Notations:

x, y	Operation sequence <i>x</i> and <i>y</i>
LCS_{xy}	Longest common subsequence between operation sequences <i>x</i> and <i>y</i>
SCS_{xy}	Shortest common supersequence between operation sequences <i>x</i> and <i>y</i>
$NOBL_x$	Number of operations of operation sequence <i>x</i> to be appended before the LCS_{xy} to form SCS_{xy}
$NOAL_x$	Number of operations of operation sequence <i>x</i> to be appended after the LCS_{xy} to form SCS_{xy}
$NOIL_x$	Number of operations of operation sequence <i>x</i> to be appended in between the LCS_{xy} to form SCS_{xy}
ξ_x	Number of bypassing moves required before the start of LCS_{xy} while producing part with the operation sequence <i>x</i> on the SCS_{xy}
ϕ_x	Number of bypassing moves required after the end of LCS_{xy} while producing part with the operation sequence <i>x</i> on the SCS_{xy}

The (2) and (3) find the minimum number of bypassing moves required before and after the LCS_{xy} while producing part with the operation sequence *x* on the SCS_{xy} . Similarly ξ_y and ϕ_y can be computed for operation sequence *y*.

$$\xi_x = \begin{cases} NOBL_y, & \text{if } (NOBL_y \leq NOBL_x) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

$$\phi_x = \begin{cases} NOAL_y, & \text{if } (NOAL_y \leq NOAL_x) \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Now the exact number of bypassing moves while producing part with the operation sequence *x* on a shortest composite supersequence of operation sequences *x* and *y* can be calculated using:

$$BPM_x = NOIL_y + \xi_x + \phi_x \quad (4)$$

Similarly the bypassing moves for producing part with operation sequence *y* on the same shortest composite supersequence is determined using:

$$BPM_y = NOIL_x + \xi_y + \phi_y \quad (5)$$

The total number of material handling movements being carried out while producing part with operation sequence *x* can simply be calculated using the bypassing moves. Thus the total number of moves while processing part *x* (TM_x) can be obtained as:

$$TM_x = BPM_x + |x| + 1 \quad (6)$$

Similarly the total number of material handling movements while producing part with operation sequence *y* is computed as:

$$TM_y = BPM_y + |y| + 1 \quad (7)$$

The number of idle machines in any layout while producing a part is an effective measure of utilization of resources; to the best knowledge of the authors, it has not

been given due consideration by the operation sequence based similarity coefficients developed so far. The number of idle machines while producing part with operation sequence x on the shortest composite supersequence of operation sequences x and y is computed using:

$$IM_x = |SCS_{xy}| - |x| \quad (8)$$

Similarly the idle machines during processing of part with operation sequence y on the same shortest composite supersequence of operation sequences x and y can be obtained as:

$$IM_y = |SCS_{xy}| - |y| \quad (9)$$

Finally after computing the exact number of minimum bypassing moves and the number of idle machines during processing each of the parts on the shortest composite supersequence of both the operation sequences, the BMIM similarity coefficient is computed as follows:

$$S_{xy} = 1 - \left[\frac{\frac{BPM_x}{2 \times |TM_x|} + \frac{BPM_y}{2 \times |TM_y|}}{\frac{IM_x}{2 \times |SCS_{xy}|} + \frac{IM_y}{2 \times |SCS_{xy}|}} \right] \quad (10)$$

The developed BMIM similarity coefficient ranges from zero to one i.e. $0 \leq S_{xy} \leq 1$ and $S_{xy} = S_{yx}$. In this similarity coefficient the percentage of bypassing moves of both the operation sequences x and y are divided by two to consider the fifty percent contribution of each operation sequence in deriving the exact similarity. Similar is the case with the percentage idle machines also.

4. ILLUSTRATION FOR COMPUTING THE BMIM SIMILARITY

For illustrating the computation of BMIM similarity coefficient consider two parts x and y with operation sequences {a w d e r t g b} and {n d t w g f h} respectively. Here the LCS_{xy} is {d t g} of length three. The length of SCS_{xy} is computed according to (1), which is 12 in this case.

Now to find the exact number of bypassing moves first of all let us consider the operations of operation sequence x which are to be inserted in the LCS_{xy} to form SCS_{xy} . The operations to be appended before the LCS_{xy} ($NOBL_x$) is two i.e. {a w}, operations to be appended after the end of LCS_{xy} ($NOAL_x$) is one i.e. {b} and the operations to be appended in between the LCS_{xy} ($NOIL_x$) is two i.e. {e r}. Similarly the $NOBL_y$ is one i.e. {n}, $NOAL_y$ is two i.e. {f h} and $NOIL_y$ is one i.e. {w}. Now according to (2) the $\xi_x = 1$ as the $NOBL_y \leq NOBL_x$ and

according to (3) $\varphi_x = 0$, as the $NOAL_y > NOAL_x$.

Similarly $\xi_y = 0$, and $\varphi_y = 1$. Now according to (4) and (5) the $BPM_x = 1+1+0$ i.e. two and $BPM_y = 2+0+1$ i.e. three. According to (6) while processing the part according to operation sequence x , the total number of moves required $TM_x = 2+8+1$ i.e. 11 and similarly according to (7), $TM_y = 3+7+1$ i.e. 11. The number of idle machines while producing part with operation sequence x can be computed using (8). Thus $IM_x = 12-8$ i.e. 4 and similarly according to (9) $IM_y = 12-7$ i.e. 5. Having all the values computed the BMIM similarity coefficient between operation sequences x and y can be calculated using (10) which in this case S_{xy} is 0.398.

5. EXISTING APPROACHES VERSUS BMIM SIMILARITY COEFFICIENT

In Table 1 four example sets of operation sequences are considered to compare the BMIM similarity coefficient with the existing coefficients. The limitations/drawbacks of various similarity/distance coefficients are highlighted in bold and underline in Table 1.

The first set is to depict the effect of precedence constraints, which are very important. According to Choobineh similarity coefficient [16], the similarity of $O_1 - O_2$ and $O_1 - O_3$ is same despite the fact that in O_3 the matched pairs are not in the sequence. The second set is taken to consider the effect of one string fully contained in other string vs string contained in the other as a substring. As in the former case only machine idleness is there and no additional material handling is required whereas in later case both the bypassing moves and machine idleness exists. The similarity/distance coefficients in [17-19] are unable to distinguish this fact and give equal similarity/distance to $O_4 - O_3$ and $O_4 - O_6$. According to [17] O_4 is more close to O_5 rather than O_3 despite the fact that O_4 is fully contained in O_3 whereas O_4 matches partially with O_5 . Irani and Huang [20] grade the similarity of $O_4 - O_3$ as 1 neglecting the effect of idle machines. The third example set is taken to depict the drawback of [18] in which first and last operations are swapped keeping rest of operations same and the similarity coefficient seems unable to recognize the similarity of matching operations. Therefore according to [18] the similarity of $O_6 - O_7$ is only 0.2 despite the three matching operations (j a c). The last example set is considered to emphasize the fact that the unmatched operations falling in between the LCS (e.g. $O_7 - O_8$) turn into both the bypassing moves and idleness of machines whereas the unmatched operations falling before the start of LCS and after the end of LCS (e.g. $O_7 - O_9$) may not depict the bypassing moves. This fact is ignored by all the existing similarity/distance coefficients including Huang similarity coefficient [21]. Choobineh shows the difference but this coefficient does not consider the precedence of matched sequences [16].

Example Set	Operation Sequence Pairs	Choobineh (1988)	Tam's distance (1990)	Ho et al. (1993)	Askin and Zhou (1998)	Irani and Huang (2000)	Huang (2003)	BMIM Similarity Coefficient
1	$O_1 - O_2$	<u>0.49</u>	0.20	<u>1.00</u>	<u>1.00</u>	0.96	0.94	0.82
	$O_1 - O_3$	<u>0.49</u>	0.60	0.50	0.50	0.60	0.49	0.64
2	$O_4 - O_5$	0.17	<u>0.43</u>	<u>0.50</u>	<u>0.50</u>	0.67	0.50	0.54
	$O_4 - O_3$	0.70	<u>0.60</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	0.94	0.70
	$O_4 - O_6$	0.20	<u>0.60</u>	<u>1.00</u>	<u>1.00</u>	0.80	0.64	0.45
3	$O_6 - O_7$	0.61	0.20	<u>0.20</u>	0.60	0.67	0.60	0.59
	$O_7 - O_8$	0.23	<u>0.49</u>	<u>0.60</u>	<u>0.60</u>	<u>0.67</u>	<u>0.60</u>	0.46
4	$O_7 - O_9$	0.42	<u>0.49</u>	<u>0.60</u>	<u>0.60</u>	<u>0.67</u>	<u>0.60</u>	0.59
	$O_7 - O_9$	0.42	<u>0.49</u>	<u>0.60</u>	<u>0.60</u>	<u>0.67</u>	<u>0.60</u>	0.59

Tab. 1. Comparison of existing similarity/dissimilarity coefficients with BMIM similarity coefficient

O1= {a c e g}, O2= {a c j e g}, O3= {e g j a c}, O4= {e g}, O5 = {e k}, O6 = {e j a c g}, O7= {g j a c e}, O8 = {g j b h e}, O9 = {b h a c e}

Note: Bold and underlines are the limitations of the similarity/distance coefficients.

As discussed above, all the existing best performing and widely used operation sequence based similarity/distance coefficients have their own limitations/drawbacks due to which the similarity between parts cannot be measured correctly. The newly developed BMIM similarity coefficient works well and gives accurate results in various situations by considering the bypassing moves and machine idleness accurately.

6. CONCLUSION

The proposed operation sequence based similarity coefficient can address most practical aspects of the part grouping problem. A more realistic approach based on longest common subsequence and shortest composite supersequence is developed to jointly consider the material flow smoothness and idleness of machines in accessing the similarity between parts. The present research leads to efficient grouping of parts in reconfigurable manufacturing system which eventually results in an improvement in the system performance measures such as machine utilization, transport utilization, throughput rate and makespan of jobs. The problems associated with operation times and alternative operation sequences may be considered in the further research.

7. REFERENCES

- [1] Mehrabi, M. G.; Ulsoy, A. G. & Koren, Y. (2000). Reconfigurable manufacturing systems: Key to future manufacturing. *Journal of Intelligent Manufacturing*, Vol. 11, No. 4, pp 403-419
- [2] Koren, et al. (1999). Reconfigurable manufacturing systems. *Annals of the CIRP*, Vol. 48, No. 2, pp 527-540
- [3] Youssef, A. M. A. & El Maraghy, H. A. (2006). Modelling and optimisation of multiple-aspect RMS configuration. *International Journal of Production Research*, Vol. 44, No. 22, pp 4929-4958
- [4] Dou, J.; Dai, X. & Meng, Z. (2009). Graph theory-based approach to optimise single product flow-line configuration of RMS. *International Journal of Advanced Manufacturing Technology*, Vol. 41, No. 9-10, pp 916-931
- [5] Goyal, K. K.; Jain, P. K. & Jain, M. (2011). Optimal configuration selection for reconfigurable manufacturing system using NSGA II and TOPSIS. *International Journal of Production Research*, (iFirst) DOI:10.1080/00207543.2011.599345
- [6] Saxena, L. K. & Jain, P. K. (2011). A model and optimisation approach for reconfigurable manufacturing system configuration design. *International Journal of Production Research*, Vol. 50, No. (12), pp 3359-3381
- [7] Abdi, M. R. & Labib, A. W. (2004). Grouping and selecting products: The design key of reconfigurable manufacturing systems (RMSs). *International Journal of Production Research*, Vol. 42, No. 3, pp 521-546
- [8] Galan, R.; Racero, J.; Eguia, I. & Garcia, J. M. (2007). A systematic approach for product families formation in reconfigurable manufacturing systems. *Robotics and Computer-Integrated Manufacturing*, Vol. 23, No. 5, Vol. 489-502
- [9] Rakesh, K.; Jain, P. K. & Mehta, N. K. (2010). A framework for simultaneous recognition of part families and operation groups for driving a reconfigurable manufacturing system. *Advances in Production Engineering & Management Journal*, Vol. 5, No. 1, pp 45-58
- [10] Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, Vol. 10, No. 8, pp 707-710
- [11] Wagner, R. A. & Fischer, M. J. (1974). The string-to-string correction problem. *Journal of the Association for Computing Machinery*, Vol. 21, No. 1, pp 168-173
- [12] Hunt, J. W. & Szymanski, T. G. (1977). A fast algorithm for computing longest subsequences. *Communications of the ACM*, Vol. 20, No. 5, 350-353
- [13] Hirschberg, D. S. (1977). Algorithms for the longest common subsequence problem. *Journal of the Association of Computing Machinery*, Vol. 24, No. 4, pp 664-675
- [14] Nakatsu, N.; Kambayashi, Y. & Yajima, S. (1982). A longest common subsequence algorithm suitable for similar text strings. *Acta Informatica*. Vol. 18, No. 2, pp 171-179
- [15] Bergroth, L.; Hakonen, H. & Raita, T. (2000). A survey of longest common subsequence algorithms. *Proceedings of 7th International Symposium on String Processing Information Retrieval (SPIRE'00, 27-29 September)*, A Coruna, Spain. New York: IEEE Computer Society, pp 39-48
- [16] Choobineh, F. (1988). A framework for the design of cellular manufacturing systems. *International Journal of Production Research*, Vol. 26, No. 7, pp 1161-1172
- [17] Tam, K. Y. (1990). An operation sequence based similarity coefficient for part families formation. *Journal of Manufacturing Systems*, Vol. 9, No. 1, pp 55-67
- [18] Ho, Y. C.; Lee, C. C. & Moodie, C. L. (1993). Two sequence-pattern, matching-based, flow analysis methods for multi-flowlines layout design. *International Journal of Production Research*, Vol. 31, No. 7, pp 1557-1578
- [19] Askin, R. G. & Zhou, M. (1998). Formation of independent flowline cells based on operation requirements and machine capabilities. *IIE Transactions*, Vol. 30, No. 4, pp 319-329
- [20] Irani, S. A. & Huang, H. (2000). Custom Design of Facility Layouts for Multiproduct Facilities Using Layout Modules. *IEEE Transactions on Robotics and Automation*, Vol. 16, No. 3, pp 259-267
- [21] Huang, H. (2003). Facility layout using layout modules. Ph.D. Dissertation, Graduate School of the Ohio State University