Cost Benefit Analysis of I*I
- Descriptive Statement of Mathematical Equations for information value and for improvement of I*I

Vijay V. Mandke

Research Leader,

Center for Information Integrity Research,

Delhi Center: B-64, Gulmohar Park, New Delhi – 110 049,

Pune Center: Flat A-2, Nikash Skies, Someshwar Wadi, Pashan, Pune-4110 08

Visit us at: centerforinformationintegrityresearch.org

Information Integrity/Integrity Information System/Management Information System
Course Lecture (s) # 50-52

2006-2007

Lecture (s) # 50-51-52
Cost Benefit Analysis of I*I
- Descriptive Statement of Mathematical Equations for information value and for improvement of I*I

OVERVIEW-1

- Cost benefit Analysis of I*I
- Gross Improvement in Information Use
- I*I Benefit Equation
- I*I Cost Equation
- Cost benefit Equation of I*I
- Criticality of Resource Commitment for Achieving Optimum I*I
- Equation for Value of Information

OVERVIEW-2

- Improvement in value of I*I due to Additional Information
- Summary
- Epilogue

COST BENEFIT ANALYSIS OF INFORMATION INTEGRITY

- Consider any information originating and processing stage ($\mathbf{S_i}$) of the *IS* view.
- Core *IS* model to which data and information are integral is modeled as **a decision process**.
- To outline the cost-benefit analysis methodology of I*I, one can consider such decision process.
  - The **decision purpose** can be to process/transform/convert data as in core *IS* to deliver information decision (by itself an information) so as **to achieve better information *use*** (for example better control for improved customer service).
- Thus the purpose of processing data/information through the core *IS* can be taken as "improvement in information *use*" **giving competitive advantage.**

GROSS IMPROVEMENT IN INFORMATION USE

- This "improvement" as variable will be a function of the information (I) being processed under the stage $\{S_i\}$ and, accordingly, it can be represented by $[\Delta IU(I)]$.
- Let IUUB(I) denote the variable giving the upper bound of information use as function of "I" (given that such function can be defined).
- Let "$\alpha(I)$" denote usefulness factor and "$\beta(I)$" usability factor.
- Both factors, functions of "I", may take values between (0,1] and, accordingly, can be seen as appropriately defined proportionality variables.
- Then, the improvement in information use at stage (S_i) is given by Equation (3).

$$\Delta IU(I)\big|_{Si} = [\alpha(I) \times \beta(I) \times IUUB(I)]\big|_{Si} \quad \text{.....Equation (3)}$$

I*I BENEFIT EQUATION

- But, reality is different as core *IS* models are complex, open and impacted by 5“C”s and they *have* errors.
- As a result there *is* a question about integrity of information “I”.
- Suppose question is about accuracy of information.
- Let “A(I)” denote the concerned integrity quotient, which takes values between (0,1].
- Then, the gross “benefit” or improvement in information use from information processing at stage (S_i) would get modified to as in Equation (4).

$$\Delta IU\ (I)\big|_{Si} = \{[\alpha(I) \times \beta(I) \times IUUB\ (I)]\big|_{Si}\} \times \{A\ (I)\big|_{Si}\} \text{.. Equation (4)}$$

I*I COST EQUATION

- Correct benefit assessment from information processing requires accounting for costs of information processing.
- What are these cost components then?
- Consistent with the individual information originating and processing nature of *IS*, it is suggested that these cost components are those of:
  - originating information "I" denoted by [$COST_{OI}$ (I)],
  - analyzing integrity quotient of A (I) denoted by [$COST_{ANAL}$ {A (I)}], and
  - the opportunity cost of analyzing A (I) denoted by [$COST_{OPPORT}$ {A (I)}]

COST BENEFIT EQUATION OF I*I

- Accordingly then the "net benefit" in the form of improvement in information use as accruing at the information processing stage (S_i) is as given in Equation (5).

$$\Delta IU(I)\big|_{Si} = [\{[\alpha(I) x \beta(I) x\ IUUB(I)]\big|_{s_i}\} x \{A(I)\big|_{s_i}\}] - [COST_{OI}(I)\big|_{Si} + COST_{ANALY}\{A(I)\}\big|_{Si} + COST_{OPPORT}\{A(I)\}\big|_{Si}] \quad \ldots\ldots \text{Equation (5)}$$

Criticality of Resource Commitment for Achieving Optimum I*I

- Assumptions:
  - Dynamic situations characterizing the information flow,
  - Considerable simplification of the query at hand
    - Assume $\alpha(I)$ and $\beta(I)$ to be given,
    - Functions IUUB(I) and A(I) having their own respective first order transients with corresponding steady state (ss) values (here of upper bound value for IUUB(I) and value equal to numerical one for A(I)), and
    - Assume all cost functions to be exponentially increasing with time,
- Then from Equation (5) the variable $\Delta IU(I)$ at the stage (S_i) under consideration *will* have a maximum value at a given time, and , among other things, *for* a given (what can be seen as an optimum, i.e., desired or, say, intended) value of integrity quotient "A".
- In other words there *is* an optimum I*I at which:
  - net increase in information use benefit is maximum (see Figure (1); Figure not to the scale);
  - achieving that I*I (implying accuracy, consistency, and reliability - if they can be quantified) is a costly process; and, to meet the demands of competitive advantage, resource commitment for achieving improved I*I, preferably optimum I*I, is critical.

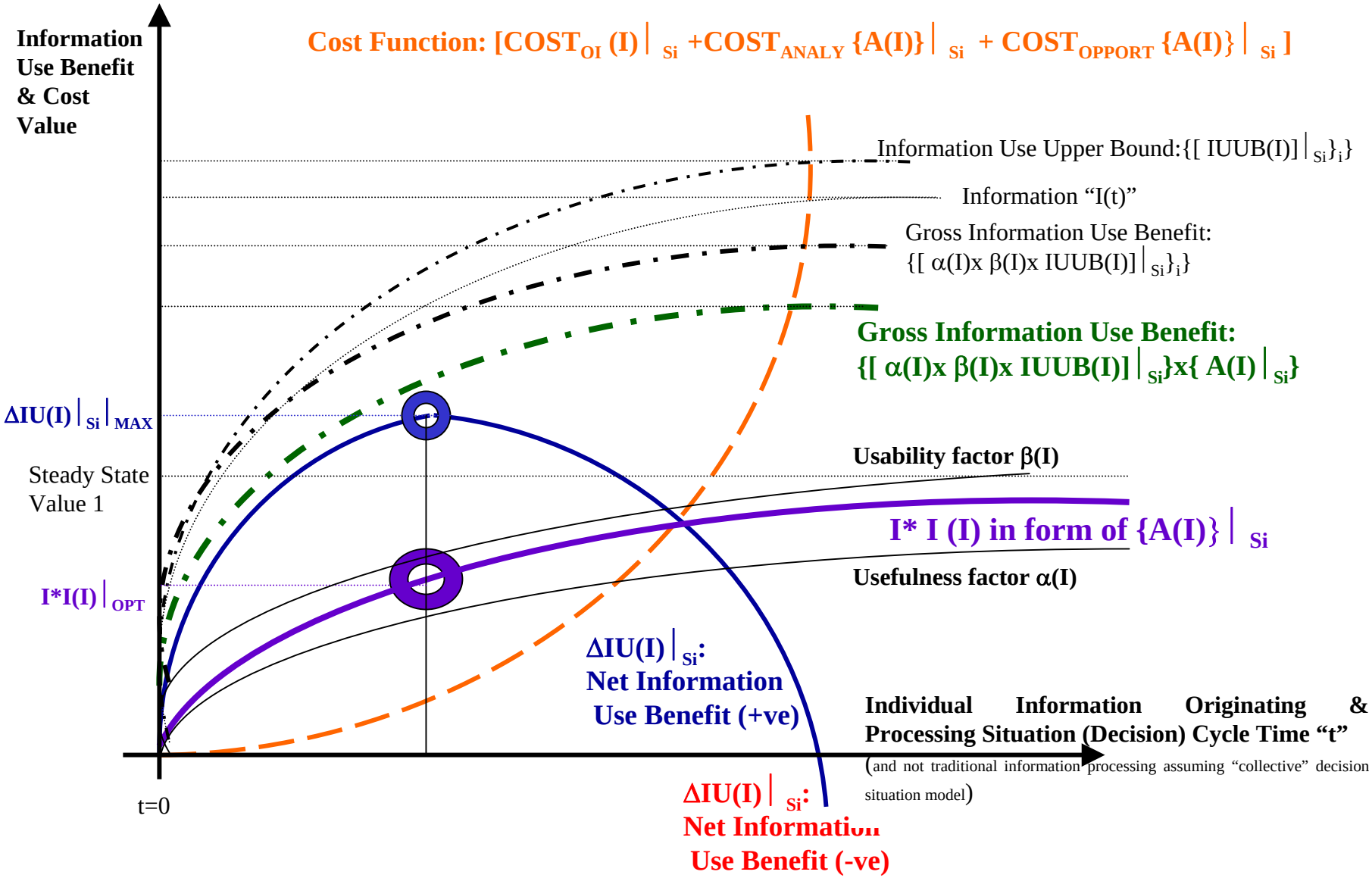

Figure (1): Cost-Benefit Analysis of Information Integrity

EQUATION FOR VALUE OF INFORMATION-1

- One could begin by formulating generalized cost-benefit equation of I*I by putting Equation (2) in Equation (5); and the same is given by Equation (6).
- $\Delta IU(I) = [\{\alpha(I) \times \beta(I) \times IUUB(I)\} \times \{A(I) \times C(I) \times R1(I) \times R2(I)\}] - [COST_{OI}(I) + \{COST_{ANAL}(A(I)) + COST_{OPPOR}(A(I))\} + \{COST_{ANAL}(C(I)) + COST_{OPPOR}(C(I))\} + \{COST_{ANAL}(R1(I)) + COST_{OPPOR}(R1(I))\} + COST_{ANAL}(R2(I)) + COST_{OPPOR}(R2(I))\]$ ..Equation (6)

EQUATION FOR VALUE OF INFORMATION-2

- We have not considered here quantification of I*I attributes. But for the purpose of investigation at hand, let us assume that [0<A(I)≤1,0<C(I)≤1,0<R1(I)≤1,0<R2(I)≤1].
- Then, it follows that [0<[I*I (I)={A(I) x C(I) x R1(I) x R2(I)}]≤1]. Information Integrity risk is then as given in Equation (7).
- Information Integrity Risk = I*I Risk =

{1–[A(I) x C(I) x R1(I) x R2(I)]}…. . Equation (7)

EQUATION FOR VALUE OF INFORMATION-3

- The question now is what maximum possible information value the *IS* modeled by Equation (6) can make use of so as to improve trustworthiness of information “I”?
- Of course, if I*I (I) has value “1”, then one can see all the determinants of I*I, that is, A(I), C(I), and R1(I) and R2(I) would be having value “1” each. In other words, there is a total confidence or trustworthiness in respect of information “I”, and, hence, it can be said there is no use getting any further information, the additional information being of zero value.
- But what if I*I (I) <1? Then, the *IS* could, according to integrity estimate provided, process additional information “I_{ADDI}” in order to improve information use; the maximum improvement in information use being possible when I*I (I) is increased to value “1”.

EQUATION FOR VALUE OF INFORMATION-4

- The question now is what maximum possible information value the *IS* modeled by Equation (6) can make use of so as to improve trustworthiness of information "I"?
- $I_{ADDIMAX} = [\alpha (I) \times \beta (I) \times IUUB(I)] - \{[\alpha (I) \times \beta (I) \times IUUB(I)] \times [A(I) \times C(I) \times R1(I) \times R2(I)]\}$

  $= [\alpha (I) \times \beta (I) \times IUUB(I)] \times \{1 - [A(I) \times C(I) \times R1(I) \times R2(I)] \} \ldots\ldots$

  Equation (8)

EQUATION FOR VALUE OF INFORMATION-5

- However, the value of the additional information "I_{ADDI}", though useful, that the *IS* actually will process may not be the maximum possible. This is because of two reasons:
  - (i) firstly, because I_{ADDI} may not have the requisite level of Information Integrity so as to increase point scales of A(I), C(I), and R1(I) and R2(I) to value of "1" each, which is necessary to bring I*I (I) to "1", and
  - (ii) secondly, because there is a need to take into account the cost components in respect of originating I_{ADDI}, and of analysing its integrity and its opportunity cost for the purpose.
- This leads in the manner of cost benefit analysis to the statement of value of information "I_{ADDI}" and the same is given in Equation (9).

$$I_{ADDI} = [I_{ADDIMAX} \times \{I*I \text{ of } I_{ADDI})\}] - [\text{Cost components concerning additional information } (I_{ADDI}) \text{ that can be usefully processed by } IS] \quad \ldots. \text{ Equation (9)}$$

EQUATION FOR VALUE OF INFORMATION- 6

- Putting Equation (8) into Equation (9) and expanding for cost components, one gets Equation (10) giving value of additional information, I_{ADDI}, in terms of Information Integrity component of information "I".
- $I_{ADDI} = [\alpha (I) \times \beta (I) \times IUUB(I) \times (1- [A(I) \times C(I) \times R1(I) \times R2(I)]) \times (A(I_{ADDI}) \times C(I_{ADDI}) \times R1(I_{ADDI}) \times R2(I_{ADDI}))] - [COST_{OI} (I_{ADDI}) + \{COST_{ANAL} (A(I_{ADDI})) + COST_{OPPOR}(A(I_{ADDI}))\} + \{COST_{ANAL} (C(I_{ADDI})) + COST_{OPPOR} (C(I_{ADDI}))\} + \{COST_{ANAL} (R1(I_{ADDI})) + COST_{OPPOR} (R1 (I_{ADDI}))\} + \{COST_{ANAL} (R2 (I_{ADDI})) + COST_{OPPOR} (R2 (I_{ADDI}))\}]$ ..Equation (10)

EQUATION FOR VALUE OF INFORMATION- 7

- Equation (10) requires calculation of I*I of I_{ADDI} — a further research investigation query, not pursued here. But, I*I of I_{ADDI} has its own attributes of accuracy, consistency, and reliability, which would need to be obtained.
- Secondly, Equation (10) is such that, apart from accounting for cost components in the manner of Equation (6), it expresses additional information value as a function of what is known about information "I" and what is actually obtained from additional information as it is processed to improve on integrity of information "I".
- And, finally, it is evident that, by controlling integrity of additional information, it is possible for IS to process higher value of useful additional information, once again reiterating the criticality of I*I (first time in respect of information "I" as in Equation (6) and, now, in respect of additional information "I_{ADDI}") for maximization of information use.

IMPROVEMENT IN VALUE OF I*I DUE TO ADDITIONAL INFORMATION -1

- With value of additional information determined, the question to be answered is the relationship between $I*I(I+I_{ADDI})$ and I*I(I).
- We have I*I(I) ={A(I) x C(I) x R1(I) x R2(I)}(Equation (2)); where I*I risk is given by [1-{A(I) x C(I) x R1(I) x R2(I)}] (Equation (7)).
- Thus, as one considers $I*I(I+I_{ADDI})$, which is given by $\{A(I+I_{ADDI})$ x $C(I+I_{ADDI})$ x$R1(I+I_{ADDI})$ x$R2(I+I_{ADDI})\}$,:
  - one is considering improvements in integrity attribute values A(I), C(I), R1(I), R2(I) of information "I" due to values of integrity attributes of I_{ADDI}; and
  - it is reasonable to suggest that each of integrity attributes of I_{ADDI} would have their respective determinants (in the manner of integrity of information on integrity).
- How does one proceed in that case? A workable approach is to consider a case when Consistency (C) and Reliability (R1) and (R2) attribute values for additional information are "1" each; that is there is complete confidence or trustworthiness in respect of Consistency and Reliability of I_{ADDI}.
- In such a situation, improvement in A(I) due to additional information "I_{ADDI}" is function of value of Accuracy attribute of I_{ADDI} denoted by $A(I_{ADDI})$.

IMPROVEMENT IN VALUE OF I*I DUE TO ADDITIONAL INFORMATION-2

- With A (I_{ADDI}) defined assuming C(I_{ADDI}), R1(I_{ADDI}) and R2(I_{ADDI}) each equal to "1", I*I attributes of information "I" would undergo improvements *only* along the Accuracy axis. This is an interesting visualization in that it encourages viewing I*I space with four dimensions in terms of: Accuracy (A) axis, Consistency (C) axis, and Reliability (R1) axis and Reliability (R2) axis. We are considering here {A (I)<1,C (I)<1,R1 (I)<1,R2 (I)<1}. Then, the information "I" has Information Integrity Risks (I*I Risk) in respect of Accuracy, Consistency, and Reliability attributes as shown in Equation (15).

(a)Accuracy (A) risk for "I" = {1- A(I)}, (b)Consistency (C) risk for "I"={1-C(I)},

(c)Reliability (R1) risk for "I"={1-R1(I)}, (d)Reliability (R2) risk for "I" ={1-R2(I)} ..Equation (15)

IMPROVEMENT IN VALUE OF I*I DUE TO ADDITIONAL INFORMATION- 3

- Due to additional information, accuracy attribute value for which is as shown in Equation (14), it is these Accuracy, Consistency, and Reliability risks of information "I" given in Equation (15), which are then to undergo reduction. And, as a result of the (extreme) case under consideration of $C(I_{ADDI})$, $R1(I_{ADDI})$, $R2(I_{ADDI})$ equal to "1", all these reductions in I*I attribute risk values for information "I", are to be effected by *only* accuracy components of I_{ADDI} . Accordingly, then, with respect to Equation (13), while maximum possible increments $\Delta A(I)$, $\Delta C(I)$, and $\Delta R1(I)$ and $\Delta R2(I)$ should facilitate removing all risks in A(I), C(I), R1(I), R2(I) raising their value to "1" each, in the case under consideration it is not so; the increments working out to be as shown in Equation (16).

$\Delta A(I) = AA(I_{ADDI}) \times \{1 - A(I)\}$, $AA(I_{ADDI})$ $\underline{\Delta}$ Accuracy component of $A(I_{ADDI})$ effecting A(I);
$\Delta C(I) = AC(I_{ADDI}) \times \{1 - C(I)\}$, $AC(I_{ADDI})$ $\underline{\Delta}$ Accuracy component of $A(I_{ADDI})$ effecting C(I);
$\Delta R1(I) = AR1(I_{ADDI}) \times \{1 - R1(I)\}$, $AR1(I_{ADDI})$ $\underline{\Delta}$ Accuracy component of $A(I_{ADDI})$ effecting R1(I);
$\Delta R2(I) = AR2(I_{ADDI}) \times \{1 - R2(I)\}$, $AR2(I_{ADDI})$ $\underline{\Delta}$ Accuracy component due to $A(I_{ADDI})$ effecting R2(I)

...Equation (16)

IMPROVEMENT IN VALUE OF I*I DUE TO ADDITIONAL INFORMATION-4

- Putting equations in Equation (16) in Equation (14), one then gets Equation (17).

$$\{ A(I_{ADDI}) \mid_{C(IADDI)=1,\ R1(IADDI)=1,\ R2(IADDI)=1} \} =$$

$$= \frac{[\{A(I)+AA(I_{ADDI}) \times \{1- A(I)\}\} \times \{ C(I) + AC(I_{ADDI}) \times \{1- C(I)\}\} \times \{ R1(I) + AR1(I_{ADDI}) \times \{1- R1(I)\}\} \times \{ R2(I) + AR2(I_{ADDI}) \times \{1- R2(I)\}\}] - [A(I) \times C(I) \times R1(I) \times R2(I)]}{[1- \{A(I) \times C(I) \times R1(I) \times R2(I)\}]}$$

$$= \frac{\text{Net change in I*I due to Additional Information}}{\text{Maximum possible change in I*I due to additional information}}$$

…… Equation (18)

IMPROVEMENT IN VALUE OF I*I DUE TO ADDITIONAL INFORMATION-5

- And rearranging the L. H. S. and R.H.S. terms, one gets Equation (18).

Net improvement in I*I (I) due to I_{ADDI} =

$$\{ A(I_{ADDI}) \mid_{C(IADDI)=1,\ R1(IADDI)=1,\ R2(IADDI)=1} \}$$

$$x\ [1- \{A(I) \times C(I) \times R1(I) \times R2(I)\}]$$

..Equation (18)

SUMMARY- 1

- Equations (17) and (18) give improvement in value of I*I in the form of $A(I_{ADDI})$ and the net improvement in I*I(I), respectively, as functions of determinants of previously accumulated information, in this case denoted by "I", and the additional information. Needless to say, similarly one will need to develop the net I*I improvement equations due to $C(I_{ADDI})$ and $R1(I_{ADDI})$ and $R2(I_{ADDI})$. This completes the investigation at hand by giving in the form of Equations (5) and (6) cost-benefit analysis of Information Integrity amenable to comparing of a set of integrity mechanisms {I*I(I)} and for selection of optimum integrity mechanism $(I*I(I)\big|_{OPT})$ by the way of maximization of net information use quantum $\Delta IU(I)$ for competitive advantage. Further, Equation (10) gives the equation for value of information, and Equations (17) and (18) give improvement in the value of Information Integrity.

SUMMARY –2

- Of course, Equation (10) has cost components for which there is a need to develop equations. Also, Equations (5) and (6), (10), (17) and (18), for calculation of respective values, are all functions of I*I attribute and I*I values previously accumulated. Thus there is the issue of developing appropriate metrics and methods for quantification of I*I attributes. In fact there can be different formulations for even Equation (2). Further, investigation here has been mainly concerned with attributes of I*I for information value, i.e., for the content of information. Information Integrity is a systems concept. This implies similarly one should develop descriptions of accuracy, consistency, reliability in respect of process integrity and system integrity and in respect of design, development, implementation, and maintenance integrity. And, for all these descriptors also there is a need to detail equations as developed in this paper. All these and many other related aspects constitute further I*I research issues.

Refer to following papers, which are uploaded onto the HSS Website for the "Reading material Link":

- 1. Vijay V. Mandke and Madhavan K. Nayar, "Cost Benefit Analysis of Information Integrity", Proceedings of the 2002 International Conference on Information Quality, MIT, Cambridge, Massachusetts, USA, 8-10 November, 2002.
- 2. Mandke Vijay V. and Nayar Madhavan K.; 'Implementing Information Integrity Technology' - Updated, Special Issue of International Journal of Informatica, June 2002.

EPILOGUE-1

- Equation (10) in the current Lecture # 27 analytically and scientifically establishes the ever increasing importance of I*I (in Lecture # 7 we first argued it qualitatively).
- Equation (6) in the current Lecture # 27 establishes requirement for optimum to control to maximize net information use in the system for achieving competitive advantage (this task we set before us in Lecture # 12, Slide # 52 [Also see Lecture # 6, Slide # 34]).

EPILOGUE-2

- Finally, Equations (17) and (18) in the current Lecture # 27 give improvement in value of I*I in the form of $A(I_{ADDI})$ and the net improvement in I*I(I), respectively, as functions of determinants of previously accumulated information, in this case denoted by “I”, and the additional information.
- With this we complete analytical treatment of *IS* model for cost benefit analysis of I*I. This is an Integrity *IS*, which will be discussed extensively through Lectures # 28-37.

THANK YOU