Conceptual Modeling of Information Quality for System Actions

Alexander Geyda
St.-Petersburg Federal Research Center
of the Russian Academy of Sciences
St. Petersburg, Russia
ageida@ias.spb.su

Abstract—The article focuses on measuring the quality of information application for actions in a system, with a focus on the pragmatic facet of this problem. The author uses various paradigms, such as system theoretic, algebraic, category theoretic, and cybernetic, along with mathematical models, to research this aspect. The results obtained are based on the author’s previous works in the field and include the concept of complex state, which consists of a measurement results substate and a higher level information substate, a set of statements about possible cause-and-effect relations in such complex states due to possible information obtained in changing conditions, and commutative diagrammatic models to explain such complex states transitions due to cause-and-effect relations.

I. INTRODUCTION

Researchers have been interested in various aspects of information usage quality, such as information value, business value of information, information technology business value, and related fields, for decades. For instance, H. Tohonen, M. Kauppinen, and T. Manisto [1], [2] note that evaluating IT business value is a challenging task, which has been on the research and practitioner agendas for more than two decades. They go on to state that value, as a multidimensional concept, is recognized as a critical factor in software and IT development, as well as decision-making within value-based software engineering.

However, despite the long-standing interest in this area, measures for using information for design, production, services, or any other human activity have not been studied in sufficient detail to predict the outcome. This is because these measures must reflect the quality of purposeful possible changes of course of actions due to information obtained and measures of their result’s fitness to the changing demands. Such models, based on mathematical formalisms, were not yet developed. This is especially true regarding predictive mathematical modeling of information use for activity and system success in changing conditions. Such modeling should be based on the description of peculiarities of the information use for activity and measures of such activity success in changing conditions. Such an approach can be considered as an extension of Batini and Scannapieco’s approach to the quality of information estimation: “we aim to investigate the relationship between the quality of information and the quality of the process output (or, simply, the process quality) that makes use of information to be produced. Since processes are made of decisions and actions, we aim in turn to relate information quality with the quality of actions and decisions that make use of information. . . We want to deepen our understanding of how the information processor, be it a human being or an automated process, can manage the fitness for use of the information consumed.” [4].

This approach is based on the concept, described by Y. Lee, R. Wang and D. Strong as: “the concept of “fitness for use” is now widely adopted in the quality literature” [5]. Decision efficiency and IQ relations, peculiarities, and available approaches for their measurement were studied. Literature review on the problem of IQ, decision efficiency estimation were provided by the aforementioned authors. A literature review of value of information estimation approaches with emphasis on fundamental, mathematical means of such estimation was provided in [1], and by the majority of other researchers based on an empirical approach. As noticed by Y. Lee, R. Wang, and D. Strong about this approach, “the disadvantage is that the
correctness or completeness of the results cannot be proven via fundamental principles”. The fitness for use is investigated by the authors [5]–[7]. As it is noticed by L. Floridi and P. Illari: “Qualitative descriptions of the meanings of words or phrases such as ‘information quality’, or ‘timeliness’ are not the same as formal metrics required to measure them, and which are needed for implementation”.

The approach suggested in the article elaborates on concepts and models proposed by the author earlier in [8]. It is based on fundamental mathematical modeling used to compute formal IQ measures based on predictive models. The contribution of the current article lies in the formalization of information use principles in the form of statements and the development of new commutative diagrams to represent information use for actions in systems. These results will allow for further formalization of information use, for example, in the form of probabilistic situation and action calculi [9] for complex states.

The objective of the article is creation of predictive models and methods which allows computing required dependencies between quality of information use characteristics and variables in various practical problems, such as information use for actions in systems success [10], digital transformation planning problems solving, new information technologies synthesis problems [11], industry 4.0 development problems [12].

The article suggests new measures that were previously proposed in [13], [14], which can be estimated using the models presented. These models use probabilistic and entropy measures to calculate formal IQ measures based on information use. The measures and models can be used to solve various problems related to information use and digital transformation, such as operations research and mathematical programming problems. The models presented in the article are concept and diagrammatic graph-theoretic models that build on the suggested concept models.

The approach presented in the article is similar to the information process modeling approach proposed by C. Batini and M. Scannapieco in [4]. However, this approach has some deficiencies, as pointed out by the authors themselves: “it does not distinguish between operational processes, which use elementary data, and decision-making processes, which use aggregated data, nor does it provide specific formalisms for them” [4]. The reason for this situation lies in the nature of information processing, which inevitably leads to purposeful changes in human action and interaction with the environment [8]. However, mathematical models that accurately capture such changes are not yet available. To improve this situation, various approaches can be used to describe changeable activity, such as the theory of functional systems [15], provided that it is operationalized with appropriate mathematical means.

II. BACKGROUND

Existing methods [16] for describing changeable activity involve building process models and traces based on log-files of processes, such as those found in healthcare and telecommunication systems. These models are typically based on Petri nets or Markov Chains. The models can then be used to enhance processes to achieve better outcomes. Enhancements may take various forms, including flexibility [17], [18] and adaptability [19]. These enhancements may be implemented through declarative, imperative, case handling, agent-based, aspect-oriented, variant approaches [20], among others.

To describe adaptive, flexible processes, change mining from change logs has been suggested [17]. The models of such processes require new methods to build and use them, including new flexible and adaptable workflow notations and models that have been developed to address possible changes in workflows. Variability patterns based on Business Process Modelling and Notation (BPMN) stereotypes were suggested [21], as well as basic adaptation patterns using BPMN 2 variants [22]. However, these approaches are not sufficient since the cause and effects of changes are not investigated and modeled as needed, especially in terms of information actions and their resulting effects. Such causes may lead to the acquisition or transmission of information of various kinds, and the list of possible kinds of information actions and their produced information is still being discussed. Examples include sensing (including monitoring, identifying, and formulating), organizing (aligning, sourcing, and combining), and restructuring (innovating, coordinating, and facilitating), or sensing, seizing, and transforming according to other sources.

Other kinds of information actions mentioned include sensing, measuring, checking, monitoring, testing, identification, organizing, research, investigation, restructuring, transforming, intelligence, and predictions. Depending on the kind of information action, various kinds of information are produced, which may cause potential changes such as necessity, reasonability, or possibility of purposeful changes. Thus, despite differences in information action kinds, which require further investigation, information actions have commonality in their potential to cause purposeful changes due to the information obtained. In the next section, the conception of IQ estimation will be discussed. Concept models will be suggested in the form of statements about information use. These models will further allow for mathematical modeling of changing system activity in changing environments, resulting in IQ measurement. These models will also allow for building graph-theoretic models that explain information application and IQ estimation, followed by building functional models and corresponding measures of IQ as measures of the success of changes due to information application in systems. The results will then be discussed and concluded.

III. CONCEPTUAL CAUSAL MODELS OF COMPLEX STATES AND ACTIONS REGARDING INFORMATION USE

In general, causal models describe the causal relations, such as functions, between variables that can describe various aspects of a system and its environment. The concept of expected value of perfect information estimation, using causal models, is related to the action taken as a result of the decision and involves measuring the maximum amount that a decision-maker is willing to pay before they are indifferent between acquiring and not acquiring information before
taking appropriate action. Unfortunately, causal models are not sufficiently aligned with systemic and cybernetic concepts of information use for system functioning. For instance, the mechanisms of information application during system functioning are not represented in the models, and characteristics of system functioning as a dependence of information use are not modeled. As a result, the capabilities of such models concerning information use for system functioning are limited. In the article, the concept of a complex state is suggested, which is further used to create models that aim to overcome the existing systemic limitations of causal models and explicitly model information use. The plan is to use the concepts and definitions suggested, elaborating the calculi of complex probabilistic states and actions regarding information use. The definitions and statements below are intended to be transformed into sorts, axioms, operations, conditions, and theorems of the calculi of complex probabilistic states and actions regarding information use to be created.

**Definition 1.** A complex state $C \in C$ of the system, including its environment(s) measurable characteristics (a set of values), and the relationships $R^{se} \subseteq S^{se} \times S^{se}$ between these values, as well as other information $I \in C$ about the system and its environment, possible modes of their functioning, and the potential outcomes of such functioning.

**Definition 2.** The information $I$ about the system and its environment (an information substrate) includes information $I^a \in I$ about the possible states $s^{se} \in S^{se}$ of the system and its environment, as well as information $I^a \in I$ about the potential outcomes of actions that could be taken based on the cause-and-effect relations $R$ between states, actions, outcomes, and measures $\Omega \in I$ defined on states, actions, and outcomes.

**Definition 3.** The information $I$ of the complex state $C \supset I$ describes the past, possible present, and future states of the system and its environment(s). Such information can be presented in the form of a known cause-and-effect structure of possible states, and known peculiarities of system and environment functioning due to the cause-and-effect structure of possible events.

**Definition 4.** Various information substrates $I$ cause various actuation (effectuation) $E(I)$ of cause-and-effect relations, including ones that are realized by humans or devices, i.e., by actuators $Ac$.

Cause-and-effect relations $R$ are initiated or carried out by humans or devices $Ac$ in order to actuate (effectuate) information in the material world. The realization of cause-and-effect relations $R$ between an information substrate $I$ and other (material) substrates $S^{se}$, including those realized by actuators, leads to the use of information for the execution of material effects.

**Definition 5.** A (sub)state of the complex state $C_i \subseteq C$ can be a type of state of the system, a state of the system and environment interaction $S^{se}$, an action $A$ state (including an information action $A^i$ state), as well as an information (sub)state $I$.

The structure of complex states $C$ makes it possible to perform sequences of various kinds of actions (parallel, sequential) and actions between the system and environment(s) $A^{se}$. For example, two substrates $(C_i, C_j)$ of a complex state make it possible to launch two actions $(A_{ij}, A_{jm})$ in various sequences with substrates $(C_i, C_j)$ as initial states of possible actions $(A_{ij}, A_{jm})$. Each initial state should have a separate information substrate to describe appropriate action sequences. An information substrate can contain, for example, descriptions of state history, current states, future states, possible causal dependencies between states, as well as observations, predictions, prescriptions, and plans. Such states may be obtained as a result of decision actions, results of information actions to obtain requirements from the environment, or results of information actions to measure correspondences of states to requirements.

**Statement 1.** Complex states $C$ change as sequences of complex states $<C_1, C_2, ..., C_n>$ associated with (possibly unknown) actions $<A_{im}, A_{jp}, ..., A_{nk}>$ in the system or its environment.

The action state can be an action beginning $C^{ab}$, an action end state $C^{ae}$, or an action during action fulfillment $C^{ae}$.

**Definition 6.** An initial complex state is a state given without the need to explicate actions that were realized for this state.

**Statement 2.** An action begins and ends with a possible complex state $S^{se}$. Action realization consists of sequences of complex states from possible ones, realized due to certain cause-and-effect relations and information (sub)state.

Each information substrate leads to possible states and cause-and-effect relations.

**Statement 3.** Actions, which differ by information (sub)states, lead to different results.

Results of the action can be modeled as a set of possible resulting states with a probability distribution defined on it, under the condition that the initial complex state was realized. The result of the action realization is a pair of complex states: the complex state of the action beginning and the complex state of the action end. The event associated with such a pair is the event of action realization.

**Definition 7.** Action beginning and end complex states can be (purely) informational, i.e., the beginning of information processing or ending state, i.e., the result of information processing. In this case, the action is informational.

**Statement 4.** The use of information for system functioning takes the form of changing action realizations depending on information (sub)states, which allows humans or devices to realize certain cause-and-effect relationships among others.
The result of information processing is an information substate, which may take the form of prescriptions for further action or descriptions in the form of diagnostic information.

**Definition 8.** An action beginning (sub)state can be a possible state of the system or environment(s) — material action start and action end.

Action-beginning complex states include information substates that allow for the changing of cause-and-effect relations in the form of various action realizations depending on information subsets. Cause-and-effect relations are always associated with an action, whether it is an information action or another kind of action. Like the state, an action can be complex, formed by possible states and causal relations between them.

**Definition 9.** Sequences of subsets of complex states can be synchronized in time or space. Action sequences synchronization can be caused by synchronization of subsets of complex states as the beginnings and ends of various actions.

**Statement 5.** Sequences of complex states and actions form a complex tree structure.

**Proof.** The sequences always begin with an initial complex state (root), with further actions potentially starting with substates and leading to different states connected by cause-and-effect relations. However, such states can never become identical due to the information substate of different actions always differing from other actions performed in parallel chains. As a result, complex states and action sequences form a tree structure.

Some sequences of actions may be considered as substructures, defined by the structure of complex states and actions between them. The case for such substructures is in the form of families of alternative stochastic networks, which are tree-structured sequences of network-shaped substructures associated with tree elements, as described in the author’s previous works [14], [23].

**Statement 6.** Complex events and actions may extend beyond the system boundary. There are two main cases of such extension: action(s) and corresponding complex events intended to obtain requirements through environment(s) and action(s) to deliver results of functioning, corresponding to requirements obtained.

Such cases can be described with transaction oriented or resource event agent architectures models. The second case serves the same goal as representing the utility node of causally structured models, used to estimate expected value of perfect information. The difference is utility value is abstract, but fitness of results transferred to environment(s) to requirements has clear interpretation in terms of functioning effects.

**Statement 7.** Quality of information use can be estimated by the complex measure, defined on the structured set of possible chains of action-complex states sequences and their results in the form of results of functioning (with use of information) fitness to requirements.

Depending on the circumstances and information obtained, including information about the requirements for the system’s functioning, various chains of complex states can be realized. Each of these sequences is characterized by a measure of results (under the condition of a certain chain of complex states realized) and requirements correspondence (“fitness”), as well as the possibility of realizing the chain of complex states. This complex measure, defined on the structured set of complex events, can be considered a measure of the quality of information use.

**Statement 8.** System actions can be divided into two main groups: (material) effects execution actions and information actions. They differ in their goals: effects execution actions are intended to obtain changes in substance and energy, while information actions are designed to obtain information changes.

**Definition 10.** Effects execution action is an action made by humans, organized by humans, or – under their control (by some technical devices) to obtain material results demanded by humans (i.e., effects).

Such effects manifested due to the exchange of energy and substance according to human’s desire, or/and under human’s control, or/and according to human’s plan performed. Exchange shall be considered in time and space and according cause-and-effects relations, which depends on information substate.

**Statement 9.** Information substate of material execution action describe possible cause-and-effects relations. Change in information state cause change in possible realizations of future action states and cause-and-effects relations changes.

Such change can be caused by additional cause-and-effects relations considered as possible (for example, if new technology became known) or by limitation of existed possibilities (for example, because of their re-evaluation).

**Statement 10.** Increasing information (i.e., increasing the amount of information available) for material execution action leads to the possibilities of new cause-and-effect relations being realized (in addition to the existing ones), or existing cause-and-effect relations being re-evaluated, thus leading to an increased possibility of achieving the required action outcomes.

During material effects execution, action information obtained (under condition it is perfect information) leads to changed to better outcomes in changed conditions.

**Statement 11.** As a result of information action, performed under condition of perfectness of information obtained for someone’s goal, the information, available for this person can not decrease.

During information action, performed for someone’s goal,
under condition of perfectness, new information obtained and older information not destroyed. It is intended to be used to alternate action for the better. Thus, possibilities to change it to the better increase. Such possibilities will especially increase if conditions of performing action and requirements for action results changed, because changing conditions and requirements more frequently require changing actions. According to the concept suggested, for material effect execution action to be executed and effects obtained, information of various kinds is required. Such information is required due to human actions’ nature, which requires operation and/or exchange of various facets of reality reflections in order to conduct action successfully. For the successful execution of actions, humans need to be sure: to begin action with required objects of the required quality, which set in required relations with other objects of interest; to begin action in certain conditions, represented as requirements to descriptions and measurements (states) of objects used in action, and to prescribe information required to act; to check states during action execution and their conformance to action prescriptions; to provide required impacts on objects and their relations during action fulfillment, according to checked states and relations of the objects used for action; to predict effects of action execution and their correspondence to requirements; to move effects received for the possible use in other actions or by other humans, through space and time. Such requirements met with information processing during action execution. They classify into three main classes: obtaining descriptive information about objects of action, about their relations, and characteristics (i.e., sensing descriptive information from objects and their relations, information processing of “in” kind, i.e., from objects); receiving, operating, and sending information from/to the information sources (i.e., processing information of various kinds without action objects affected); using the information to provide required impacts on objects, and their relations during action fulfillment, described by this information (i.e., actuating prescriptive information, kind of information processing of “out” kind, i.e., to objects).

Subsequently, three kinds of information processing for the execution of action effects are distinguished. Furthermore, three kinds of information processing per se, not classified as ‘in’ or ‘out’ types (i.e., without the participation of objects used for the execution of effects), are also distinguished. Such information processing, without the participation of objects used for material effects execution, can be classified into three kinds of information: descriptive (hindsight, answering ‘what happened’), prescriptive (insight, answering ‘how to make it happen’), and predictive (foresight, answering ‘how do we make it happen’) information, and combinations of them.

Predictive information processing differs because it does not necessarily produce information about particular action objects used to effect execution. It produces higher-level information. Predictive information is general information about why and how effects manifest, not only in specific action execution cases with particular objects and information used for them. There may also be other higher-level types of information processing that do not necessarily reflect the effects of actions. This higher-level information processing is a kind of information processing used to obtain explanations, rules, peculiarities, and predictions about the results of object use, general laws of nature functioning, their descriptions, and so on. It is also used to form predictions about different action fulfillment, formation rules for different human requirements, and probably other levels of explanations. These explanations consist of knowledge about the different ways humans and nature activities form and other phenomena details due to the actions of humans and nature. This article does not consider processing information of these higher-level kinds yet; it is subject to future research.

As the authors noted previously, information processing during effects execution actions should be distinguished in time (before, during, after actual effects execution) and in space. For this reason, various kinds of synchronization models are provided, like cyclograms models for time synchronization among actions or technological routes for modeling moving material entities and energy in the space.

Statements formulated intended for the purpose of future definition of formal calculi of complex states and actions, including actions to process information substates.

To describe such calculi properties, diagrammatic models of complex states and actions changes due to information use were constructed. They are described in the following section.

IV. DIAGRAMMATIC MODELS OF COMPLEX STATES AND ACTIONS CHANGES REGARDING INFORMATION USE

To represent possible state changes due to information application, and information actions of various kinds, as described in the previous section, commutative diagram notation was developed. An example of such a diagram can be seen in Fig. 1. These diagrammatic models can be used to create possible sequences of state and action changes depending on the information used, which can then be used to estimate the quality of information use based on the characteristics of the actions and IT used.

In the figure:

- $S^{ES}_{ES}$ is the complex state of the system and its environment before information use.
- $I^{ES}_{ES}$ is the information action used to check the environment and system states.
- $I^{ES}_{ES}(S_i)$ is the information that describes the checked environment and system states.
- $i^P(n)$ is the information action used to produce prescriptions for further action based on the information obtained from state’s check.
- $I^P_{ES}(S_i)$ is the information that describes the chosen prescriptions to fulfill further action among possible ones.
- $i^E(n)$ is the information action used to deliver prescriptions for further fulfillment by an actuator, whether human or device.
- $S^{ES}_{ES}(n)$ is the state that describes the possible prescriptions to fulfill further action before information is applied.
- $s^{ES}(n)$ is the substate that describes the chosen prescriptions to fulfill further action obtained after information is applied.
The case for information obtained separately from the system and its environment is represented in Fig. 2.

For comparison, the action to prepare prescriptions without taking into account the information obtained about the current states of the system and its environment is illustrated in Fig. 3.

As a result, (sub)sets, describing possible further actions, remains unchanged (because no information obtained to change them).

**Definition 11.** Semi-commuting diagrams consists of objects, which are complex sets, determined earlier, with probability distributions, defined on σ-algebras of subsets of complex sets, and arrows which are mappings of objects.

The diagram is semi-commuting because its information application arrows can be maps to subset or maps to upper sets, or maps to sets with changed measure defined on the complex sets, as a result of information application.

**Statement 12.** Information application in the form of mapping to subset of a complex set consists of limiting possible outcomes, represented by subsets, to outcomes which suites better to obtained information.

The diagram in Fig.1 is this case, and it is partially commuting in that sense, state $s^{n}(n) \subset S^{ES}(n)$, i.e. due to information application actions sequences variability of states for further action beginning becoming less, because some sets are no longer considered as realizable. But such reduction in variability (assuming perfect information) leads to a higher likelihood of future results meeting the requirements, as better states were chosen as a result of the information operation. In other words, information application through mapping to a subset of a complex set allows for a more focused and informed decision-making process, ultimately improving the quality of the subsequent actions.

**Statement 13.** Information application can be realized in the form of mapping to the upper set of a complex set consists of enlarging the set of possible outcomes, represented by subsets. As a result, outcomes which will suites better to information obtained, can be found in complement of the initial set to upper set.

An example of such application may be the modernization of the technological base, when sets of possible of actions enlarge due to information obtained. Illustration of such case can be found in Fig. 4

**Statement 14.** Information application can be realized in the form of mapping to the upper set of a complex set, which consists of enlarging the set of possible outcomes represented
by subsets. As a result, outcomes that better suit the obtained information can be found in the complement of the initial set to the upper set.

Sequences of semi-commuting paths on diagrams may serve as a basis to estimate sequences of states and their measures for estimation of the information application based on measures of results correspondence to requirements and measures of complex states possibilities.

**Statement 15.** Probability distributions, defined on complex subsets, associated with diagrams objects, according to chains of objects, defined by arrows, may serve as a basis to build complex measures to estimate quality of information application.

Examples of such measure computation can be found in [13].

**V. Consequences and Applicability of Results.**

The statements and diagrammatic models presented serve as the basis for creating complex probabilistic states models with regard to information use for adaptive actions. These models used to allow calculus of complex states and actions with regard to information use. Such calculus could enable the automatic generation of possible sequences of complex states and events regarding information use, which is currently done manually.

Within the calculus to be created, each diagrammatic model presented uses as many combinations of input information as possibly exist, acting as a probabilistic generator of possible strings of complex events, depending on information realized. These strings can be considered as sequences of actions results obtained due to the presented diagram’s use for sequences generation. The strings obtained from realizations are saved in a tree-like or table structure.

Such structure further allows evaluating suggested [24] measures of results correspondence to changing conditions, measures of results actualization and measures of variability of such complex states chains. These measures used to construct functional dependencies of complex measure of activity quality regarding information use variables.

These functional dependencies between quality and variables allow solving variety of problems related to information use for actions in systems, digital transformation problems, new information technologies construction.

In the example considered, sequences for evaluation of quality measures are modeled as branches of the tree of possible complex states changes and then saved as rows in a data table that describes possible sequences of information use for action in the system, depending on various conditions. Measures are computed for each row, which represents a tree branch. Each branch’s measure of result correspondence (fitness) is calculated, along with the branch’s possibility of being actualized. A detailed description of the measures’ computations and the example table formation details can be found in [24].

The table represents characteristics of possible chains of complex events and states in the probabilistic space of possible chains of complex events and states. Characteristics of such possible chains are computed in the example by the table rows, which allows computation parallelization.

Complex measures are computed for the table as a whole. The complex measure for the tree is calculated as probabilistic mixture measures and mixture entropy measures [25].

An example table with the data used and probabilistic mixture measures computed is shown in Table I.

**VI. Conclusion**

The article proposes methods to model information application for actions in systems. It suggests using diagrammatic models based on complex states and actions, which are linked through cause-and-effect relationships. Complex states consist of possible alternative substates with defined probability distributions, and subsets of these states can be activated through actions. The objective of the article was to build, based on models and methods suggested, required dependencies between quality of information use characteristics and variables. Such dependencies required to solve a variety
of problems related to information use for actions in systems, digital transformation problems, new information technologies synthesis. Partially commuting diagrammatic models of complex states and actions were used for this purpose. They allow generating sets of possible sequences of complex states due to information application, resulting in a tree of complex states. This tree was used to build a table of possible characteristics of chains of complex states. The suggested models were used to estimate system functioning and the quality of information use, depending on the parameters and variables of the information actions and corresponding technologies used by the system. The results obtained are intended to serve as the basis for creating adaptive actions and complex probabilistic states calculus regarding information use.

REFERENCES


