HIGH-FIDELITY MATHEMATICAL MODELS OF SOCIAL SYSTEMS

Joe Jeffrey[1], Northern Illinois University, DeKalb, IL

ABSTRACT

We present a mathematical formalism, Entity Specifications, with sufficient rigor and expressive power to formalize a wide range of models of social systems. Entity Specifications are a combination of a rigorous formulation of the intuition of a frame and formal representation of relationships among constituents, as in mathematical logic. Entity Specifications are then used to formalize a particular conceptual framework that has been used as the basis for a number of computer systems, Communities and Intentional Action. This model articulates a broad range of social system phenomena, from individual actions to actions of groups at all levels, encompassing all types of phenomena actually encountered in human systems: biological, psychological, economic, sociological, political, and cultural. The formalism is then used to develop new mathematical formulations of concepts with broad applicability in the social sciences: complexity, similarity, and rate of change and rate of complexity change in social systems. The similarity measure is illustrated with an example in which the similarity between two pairs of intuitively similar families is calculated.

This paper presents a mathematical formalism for building high-fidelity models of the structure and dynamics of social systems. By "high-fidelity" we mean accurate representation of the entire range of the situations, processes, and events in the system, at every level of detail. The central intuition of the formalism is that a "thing" – an object, process, or state of affairs – is specified by giving a formal name, the logically necessary immediate constituents, and the relations between the constituents, all constituents and relationships also specified by formal name, as in mathematical logic. Constituents themselves may be further elaborated in the same way. The formalism is designed to handle incomplete knowledge and does not require reduction to atomic elements.

ENTITY SPECIFICATIONS

We need to be able to formally describe three kinds of "thing": objects (structures), processes (mechanisms), and states of affairs. Entities have parts – immediate constituents – that may be objects, processes, or states of affairs. An entity is described by giving its name and a description; the description consists of the names of the constituents and their type (object, process, or state of affairs), and their relationships.

Definition An *entity specification* (*ES*) consists of an ordered pair (N, D), where:

- N is the (formal) name of the entity including, optionally, a list of alternate names and/or a numerical ID.
- D is the set of *paradigms*, the major varieties or descriptions of the entity. In social systems, these are often structures or processes with little in common other than being

[1] H. Joel Jeffrey, Dept. of Computer Science, Northern Illinois University, DeKalb, IL 60115. Email: jeffrey@cs.niu.edu.

recognizable as varieties of the same thing. For example, in Western society celebrating a wedding anniversary has several paradigms: dining out, taking a cruise, buying a gift, etc.

Each paradigm of D is an ordered triple (C, R, E), where:

- C = $\{(C_i, T_i)\}$, in which C_i is the constituent and T_i is the constituent's classification, an element of the set {P, O, S}, representing "process,""object," or "state of affairs."
- R is the set of n-ary *relationships* that must hold between the named constituents. Any relationship may be named, not only those definable in terms of physical, mathematical, or computable quantities. (Equations are formal names.)
- The constituents and their relationships specify the structure of the entity. Additional information specifies particular instances of the entity. Identifying an actual instance requires the specification of which actual "things" (processes, objects, and states of affairs) that fill the roles named by the constituents. This information we term the *eligibilities* E for the entity: a set of ordered triples (c, i, r), in which
    - c is the name of the constituent;
    - i is the name of the individual;
    - r is the rule, or condition, under which i takes the role of c in this object.

ESs are a formalization and unification of the representation formats developed by P. G Ossorio[2], as a means of specifying objects, states of affairs, and processes at any level of detail.

A state of affairs' constituents may be any set of objects, processes, and other, smaller, states of affairs, in various relationships, i.e., any set C = $\{(C_i, T_i)\}$ and any set R, as defined above.

Processes

Processes are multi-step changes in objects and their relationships. Processes may occur in many ways, i.e., combinations of the steps. Therefore the $\{(C_i, T_i)\}$ for a process include:

1. Two constituents, the before-state and after-state.
2. A subset identifying stages, i.e., in which T_i = P. Some stages may be accomplished via two or more alternatives; these alternatives are included in this subset.
3. A subset identifying the objects, i.e., T_j = O.
4. A subset identifying the versions of the process. Each of these constituents C_k is a state of affairs, i.e., T_k = S, and the constituents of C_k are stages.

Relationships between stages specify the time relationships between them: sequential, parallel, overlapping, interspersed, etc.

Objects

Objects have only object constituents, and in that sense are simpler than entities in general or processes; each constituent of an object of Type O.

[2] P. G. Ossorio, *"What ActuallyHappens"*, University of South Carolina Press, Columbia, SC, 1978. Republished by Descriptive Psychology Press (www.descriptivepsychologypress.com), Ann Arbor, MI, 48104, in 2005.

Relationship to Frames

ESs are similar to frames, but differ in two important ways. First, ESs are considerably more rigorously defined. The constituents of an ES must be logically necessary for the thing being described to be what it is. By contrast, a frame simply represents "things commonly found together[3]." No distinction is made between things that are necessary to the definition of thing and those that are merely commonly present.

Second, frames do not include specification of the $R_1...R_m$ between constituents, although some frame-based systems allow specification of relationships. (Interestingly, while clearly a refinement of frames, Ossorio's work predates the introduction of frames by several years[4].)

"Incomplete" Descriptions

Most real social systems are too complex to be specified completely, i.e., at the level of actions by individual persons. ESs are designed for handling incomplete specifications: a description and its elaborations are used in analysis or simulation by modeling the named entities and their interactions at that level. An ES set is not like a computer program, which is incomplete and not executable until all functions have been written in executable primitive statements.

For example, the state of affairs identified by the English sentence, "The rise of inflation in 1920's Germany led to the rise of National Socialism," identifies an entity (a state of affairs) with two constituents: "the rise of inflation in 1920's Germany," "National Socialism," and the relationship "led to the rise of." Each of these constituents has further Descriptions in terms of constituents and relationships, and a set of descriptions down to the level of the famous image of a woman with a wheelbarrow of Deutsche marks to buy a loaf of bread would be enormous. It is not, however, necessary for simulation or analytical treatment.

DESCRIBING SOCIAL SYSTEMS

Social systems, distinguished by having human "components," require a conceptualization encompassing the specifically human aspects of action in a social context. The most comprehensive and systematic conceptualizations of this subject matter of which we are aware are the Community formulation, due to Putman[5] and Intentional Action, due to Ossorio[6]. They have been successfully used to build formal models of a number of social systems, and

---

[3] Minsky M (1975), "A Framework for Representing Knowledge", in Winston P, ed., *The Psychology of Computer Vision*. New York: McGraw-Hill, NY.
[4] P. G. Ossorio, *State of Affairs Systems: Theory and Technique for Automatic Fact Analysis,* RADC-TR-71-102, Rome Air Development Center, 1971.
[5] A. O. Putman, "Communities," in *Advances in Descriptive Psychology,* V. I, K. E. Davis ed., JAI Press, Greenwich, CT, 1981.
[6] P. G. Ossorio, *The Behavior of Persons,* Descriptive Psychology Press, Ann Arbor, MI, 48104, 2006.

software based on the models[7,8]. We illustrate using ESs to formalize social systems by formalizing Communities and Intentional Action. However, entity specifications are not particular to this conceptualization; formalizations of any model, no matter how abstract, physicalistic, or even fanciful, may be given.

A **Community** description is a 4-tuple <M, P, Cp, S>, where

- M (members) denotes all actual individuals in the community.
- P (Practices) denotes the set of social practices of the community. Practices encompass everything that a member of that community can do, *as* a member of that community.
- Cp (Choice principles) denotes the set of values or priorities specific to the community. The principles govern which practice is carried out and how, playing a large role in determining what actually occurs and how it occurs in the social system.
- S (Statuses) denotes the recognizable positions in the Community, whether formal and explicit or informal and implicit. “President,” “Senator,” “husband,” “child,” suicide bomber,” “respected leader,” “doctor,” “farmer,” etc., are examples. Each position has associated with it one or more *intrinsic practices*, practices a member engages in simply because they are in that position.

A practice is specified by giving a s*ocial practice description*, a quintuple <W, K, Kh, P, PC>, where

- W (want) denotes the goal.
- K (know) denotes the facts and concepts necessary for this action.
- Kh (know-how) denotes the skills needed to carry out the practice effectively.
- P (process) denotes the procedural aspect of the practice.
- PC (personal characteristics) denotes any relevant attitudes, traits, or abilities.

Finally, to specify the details of what happens in the system, the process parameter itself must be expanded with process entity description: the stages, constraints, eligibilities, and versions of the process.

What actually occurs is an *instantiated version* of the practice: a particular set of stages, with specific individuals filling the logical roles E, much as an actual production of *Hamlet* consists of the sequence of Act of the play, with actual persons and objects filling the roles of the characters and props.

To engage in an action is to engage in the practice of a community. The operation of the social system consists of members engaging in practices, in accordance with the choice principles of the community.

---

[7] “MENTOR: Replicating the Functioning of an Organization", in *Advances in Descriptive Psychology*, Vol. III, pp. 243-270, K. E. Davis, ed., JAI Press, Greenwich, Connecticut, 1983.

[8] Jeffrey, H. J., Schmid, T, Zeiger, H.P, and Putman, A. O., 1989, "LDS/UCC: Intelligent Control of the Loan Documentation Process" *Proceedings of the Second International Conference on Industrial & Engineering Applications of Artificial Intelligence and Expert Systems*, University of Tennessee Space Institute, Tullahoma, Tennessee, June, 1989, ACM Press, pp. 573-591.

ENTITY DESCRIPTIONS OF COMMUNITIES AND PRACTICES

Combining entity specifications with community descriptions gives a formal model of a social system. A community is an entity, with object, process, or event constituents: Members, a set of object names; Practices, a set of entity specications; Choice principles, a set of state of affairs entities; Statuses, a set of states of affairs entities.

Communities and intentional action, stated in Entity Specification form, provide a mathematical representation of human behavior in the human context, at any level of detail. A social system is a formal entity consisting of (formal) immediate constituents, with n-ary relationships between them, and elaboration of constituent entities, at any level, via entity specifications.

As noted above, the ES formalism may be used to formalize any model of human behavior; the community and practice model formulations are but one, albeit the most comprehensive and systematic we know of.

MEASURING SIMILARITY AND COMPLEXITY

Having a formalization of social systems allows us to give new mathematical formulations of the concepts of complexity of a social system and similarity of two social systems.

We first define the *structural complexity* of a social system A, with with N constituents $A_1, ..., A_N$ and K relationships, as:

$$SC(A) = \sqrt{N^2 + K^2 + \varepsilon \bullet \sum_{i=1}^{NA} SC(A_i)^2}$$

ε is an experimentally-determined multiplier modulating the impact of complexity of constituents, sub-constituents, etc. (Preliminary work indicates a value of approximately 0.7 for ε.)

We can now mathematically define the degree of similarity between any two entities based on their constituents and relationships. The definition is designed to correspond to the following intuitions:

1. The measure should take into account differences in attributes of the entities themselves.
2. The measure should take into account similarity of structure. Structure is formalized by relationships among constituents, in two ways: a) differences in the attributes of the constituents of A and B; b) if A and B have the same relationships among their respective constituents, but to different degrees, similarity should reflect the difference in degree; c) if A and B have different relationships, they should be less similar.

3. When the constituents of A and B themselves have ESs, the measure should recursively include the structural similarity of the constituents.

Accordingly, the distance between two entities is defined as follows:

1. Let A and B be any two entities, the properties of A and B be $p_1, ..., p_M$, and entity specifications comprised of constituents $A_1, ..., A_{NA}$ and relationships $r_1, ..., r_K$, and $B_1, ..., B_{NB}$ with the L relationships $r_{K+1}, ..., r_{K+L}$.

   The constituents $A_1, ..., A_m$ with relationship r_j are ordered m-tuples. Denote the number of A-tuples by NAT, and the number of B-tuples by NBT.

2. Let the value of property i of a constituent be represented by $p_i()$, and $r_i(t)$ denote the value of the ordered tuple t of A- or B-constituents satisfying relationship r_k. For example, a strong love relationship between family members A_1 and A_2 is represented by loves(A_1, A_2) = 0.9 (on a 0 to 1 scale).

   Let P denote the matrix with M columns and NA+NB rows, whose values are the values of each property p_i.

P:

	p_1	...	p_M
A_1			
...			
A_{NA}			
B_1			
...			
B_{NB}			

   The *property distance* between A_i and B_j is given by

$$PD(A, B) = \sqrt{\sum_{i=1}^{M} (p_k(A_i) - p_k(B_j))^2}$$

   Let R be the matrix with K+L columns and NAT+NBT rows whose entries are the values $r_k(t)$. If a constituent does not have property p_i, or a tuple does not have relationship r_k, leave the corresponding entry of the matrix blank.

R:

	r_1	...	r_K	r_{K+1}	...	r_{K+L}
A-$tuple_1$						
...						
A-$tuple_{NAT}$						
B-$tuple_1$						
...						
B-$tuple_{NBT}$						

3. If any column of P or R contains a value < 0, re-scale the values of the column by adding the absolute value of the minimum value of the column to each value in it.

4. Normalize the values of P to the range 1 to 10, by setting

   $p_i(A_j) = 10 * (p_i(A_j) + 1) / pmax_i$), where $pmax_i$ is the maximum value of column i.

5. Set each empty entry of P to 0.

   The values of the property matrix P are now between 0 and 10, 0 indicating the component does not have the property of that column.

6. Similarly, normalize the values of R, the matrix with K+L columns representing relationship values for constituents of A and B, to the range 0 to 10.

7. When A and B have constituents, the similarity between A and B must include similarity of their constituents. That calculation is affected by the order of the constituents. For example, suppose A and B are organizations, and A has a large and complex marketing department and a small, simple shipping department, while B has a large and complex shipping department and small, simple marketing department. The calculated similarity between A and B will be quite different, depending on whether the two marketing departments and two shipping departments are compared, or whether A's marketing department is compared to B's shipping department. The constituents of each must be re-ordered so that the distance comparison has a consistent basis.

   Therefore, re-order the constituents of A and of B, from maximum SC (as defined above) to minimum.

8. The **distance between two entities A and B** is comprised of two components, the property distance and the structural distance:

   $$d(A, B) = \sqrt{PD(A, B)^2 + SD(A, B)^2}$$

   The *structural distance* SD(A, B) is defined recursively as follows:

   Let MC = max(NA, NB) and MT = max(NAT, NBT).

Then if both A and B have Descriptions, i.e., specified constituents and relationships, we define the structural distance SD as

$$SD(A, B) = \sqrt{(NA-NB)^2 + \sum_{i=1}^{MC} PD(A_i, B_{\kappa(i)})^2 + \sum_{j=1}^{MT} \sum_{i=1}^{L} (r_i(ta_j) - r_i(tb_{\kappa(j)}))^2 + \delta \bullet \sum_{i=1}^{MC} SD(A_i, B_i)^2}$$

where $B_{\kappa(i)}$ denotes the B-constituent closest to A_i, using Euclidean distance, and

$tb_{\kappa(j)}$ denotes the B-tuple closest to ta_j, using Euclidean distance between tuples.

If A NAT > NBT, $r_i(tb_j) = 0$ for NBT < j <= NAT, and similarly if NBT > NAT.

If NA > NB, then $d(A_i, B_i) = SC(A_i,)$, for NB < i <= NA, and similarly if NB > NA.

If either A or B have no Description, SD(A, B) = 0.

δ is an experimentally-determined discount factor reflecting the relative importance of the distance between constituents of A and B. (As with ε, preliminary work indicates a value of approximately 0.7 for δ.)

$PD(A_i, B_i)$ measures similarity of properties of each pair of constituents;

$\sum (r_i(ta_j) - r_i(tb_j))^2$ measures how much the constituents of A and B differ on relationship r_i; and the sum $\sum_{j=1}^{MT} \sum_{i=1}^{L} (r_i(ta_j) - r_i(tb_{\kappa(j)}))^2$ measures the total difference in structure between A and B, as articulated by the relationships r_i, $1 \leq i \leq L$.

The distance measure d(A, B) has the following properties:

- d(A, B) = 0 if A and B are the same except for differing only in names of constituents and relationships (mathematically, are isomorphic).
- The distance increases as the properties of A and B, the number of their constituents, the properties of the constituents, the structure of A and B, and the substructures of A and B diverge.

As with any mathematical definition intended to capture an intuition, this formulation must validated experimentally. This work is in progress.

Example: Structural Similarity of Two Families

Family A consists of a mother, father, and two children. The mother and father are married, and love each other. Both parents love both children; the children love each other. However, the children also compete with each other for success in school.

Family B consists of a mother, father, and three children. The mother and father are married. Both parents love all the children. The two younger children love each other, but both resent the eldest and compete with her for each parent's affection. The eldest child also has a significant responsibility in caring for the younger children

We suppose that the member attributes of interest in this case are age and health of the family members.

	Age	Health
M_A	40	0.8
F_A	42	0.7
AC_1	12	1.0
AC_2	10	1.0
M_B	35	0.9
F_B	36	0.8
BC_1	8	1.0
BC_2	6	1.0
BC_3	14	1.0

Table 1: P matrix for Families A and B

	Rom.. Love	Par. love	Sib. love	Aca. comp.	Resent	Affec. comp.	Care-taker
(M_A, F_A)	1.0						
(F_A, M_A)	1.0						
(M_A, AC_1)		1.0					
(M_A, AC_2)		1.0					
(F_A, AC_1)		1.0					
(F_A, AC_2)		1.0					
$(AC_1\ AC_2)$			1.0	1.0			
(AC_2, AC_1)			1.0	1.0			
(M_B, F_B)	1.0						
(F_B, M_B)	1.0						
(M_B, BC_1)		1.0					
(M_B, BC_2)		1.0					
(M_B, BC_3)		1.0					
(F_B, BC_1)		1.0					
(F_B, BC_2)		1.0					
(F_B, BC_3)		1.0					
(BC_1, BC_2)			1.0				
(BC_2, BC_1)			1.0				
(BC_1, BC_3)					1.0		
(BC_2, BC_3)					1.0		
(BC_1, BC_3)						1.0	
(BC_2, BC_3)						1.0	
(BC_3, BC_1)							1.0
(BC_3, BC_2)							1.0

Table 2: R matrix for Families A and B

The normalized values of the properties of the families A and B are (8.2, 10, 10) and (10, 10, 10), so PD(A, B) = $\sqrt{1.8^2 + (10\text{-}10)^2 + (10\text{-}10)^2}$ = 1.8.

$(NA–NB)^2 = (5\text{-}3)^2 = 4$.

The normalized property matrix P, with rows re-ordered so that the pairs A_i, $B_{\kappa(i)}$ are adjacent, is

	Age	Health
M_A	9.5	8
F_B	8.6	8
F_A	10.0	7
M_B	8.3	9
AC_1	2.9	10
BC_3	3.3	10
AC_2	2.4	10
BC_1	1.9	10
BC_2	1.4	10

Table 3: Normalized P for Families A and B

and the constituent property distance $\sum_{i=1}^{5} PD(A_i, B_{\kappa(i)})^2 = 0.81+2.89+0.16+0.25+1.96$

$= 6.07$

The normalized relationship matrix R, with rows re-ordered so that the pairs ta_j and the nearest tuple $tb_{\kappa(j)}$ are adjacent, is

	Rom.. Love	Par. love	Sib. love	Aca. comp.	Resent	Affec. comp.	Care-taker
(M_A, F_A)	10	0	0	0	0	0	0
(M_B, F_B)	10	0	0	0	0	0	0
(F_A, M_A)	10	0	0	0	0	0	0
(F_B, M_B)	10	0	0	0	0	0	0
(M_A, AC_1)	0	10	0	0	0	0	0
(M_B, BC_1)	0	10	0	0	0	0	0
(M_A, AC_2)	0	10	0	0	0	0	0
(M_B, BC_2)	0	10	0	0	0	0	0
(F_A, AC_1)	0	10	0	0	0	0	0
(M_B, BC_3)	0	10	0	0	0	0	0
(F_A, AC_2)	0	10	0	0	0	0	0
(F_B, BC_1)	0	10	0	0	0	0	0
$(AC_1\ AC_2)$	0	0	10	10	0	0	0
(BC_1, BC_2)	0	0	10	0	0	0	0
(AC_2, AC_1)	0	0	10	10	0	0	0
(BC_2, BC_1)	0	0	10	0	0	0	0
(F_B, BC_2)	0	10	0	0	0	0	0
(BC_1, BC_3)	0	0	0	0	10	0	0
(F_B, BC_3)	0	10	0	0	0	0	0
(BC_2, BC_3)	0	0	0	0	10	0	0
(BC_1, BC_3)	0	0	0	0	0	10	0
(BC_2, BC_3)	0	0	0	0	0	10	0
(BC_3, BC_1)	0	0	0	0	0	0	10
(BC_3, BC_2)	0	0	0	0	0	0	10

Table 4: Normalized R for Families A and B

The structural-difference sum $\sum_{j=1}^{MT} \sum_{i=1}^{L} (r_i(ta_j) - r_i(tb_{\kappa(j)}))^2 = \sqrt{6 \bullet 0^2 + 2 \bullet 10^2 + 8 \bullet 10^2} = 31.6$

In this example, the immediate constituents are individual persons. Customarily one considers persons to be indivisible, so in this case $SD(A_i, B_i) = 0$. (In a case in which it is considered appropriate to model parts of a person, such as aspects of personality, $SD(A_i, B_i)$ may be non-zero.)

Thus $SD(A, B) = \sqrt{4 + 6.07 + 31.6}$ and $d(A, B) = \sqrt{1.8^2 + 4 + 6.07 + 31.6} = 6.7$

Consider now the distance between A and B', a family identical to B except that $B'C_1$ and $B'C_2$ did not resent and compete for affection with $B'C_3$. The rows of R representing those relationships are now missing, so the structural-difference sum is 24.5, and d(A, B') = 6.1.

Rate of Social Change

Thus, we can now define rates of change in a social system:

The **rate of social change**, as a system goes from S_1 to S_2, in time Δt, is $\frac{d(S_1) - d(S_2)}{\Delta t}$

and the **rate of social complexity change** of S is $\frac{SC(S_1) - SC(S_2)}{\Delta t}$.

REFERENCES

Jeffrey, H. J., 1983, “MENTOR: Replicating the Functioning of an Organization", in *Advances in Descriptive Psychology*, Vol. III, pp. 243-270, K. E. Davis, ed., JAI Press, Greenwich, Connecticut.

Jeffrey, H. J., Schmid, T, Zeiger, H.P, and Putman, A. O., 1989, "LDS/UCC: Intelligent Control of the Loan Documentation Process," *Proceedings of the Second International Conference on Industrial & Engineering Applications of Artificial Intelligence and Expert Systems*, University of Tennessee Space Institute, Tullahoma, Tennessee, June, 1989, ACM Press, pp. 573-591.

Minsky, M, 1975, “A Framework for Representing Knowledge”, in Winston P, ed., *The Psychology of Computer Vision*. New York: McGraw-Hill, NY.

Ossorio, P. G., 2006, *The Behavior of Persons*, Descriptive Psychology Press, Ann Arbor, MI 48104 (www.descriptivepsychologypress.com).

Ossorio, P. G., 1971, *State of Affairs Systems: Theory and Technique for Automatic Fact Analysis,* RADC-TR-71-102, Rome Air Development Center.

Ossorio, P. G., 2005, “*What Actually Happens”*, University of South Carolina Press, Columbia, SC, 1978. Republished in by Descriptive Psychology Press, Ann Arbor, MI 48104. (www.descriptivepsychologypress.com).

Putman, A. O., 1981, “Communities,” in *Advances in Descriptive Psychology*, V. I, K. E. Davis ed., JAI Press, Greenwich, CT.