Measuring relational complexity in oddity discrimination tasks

Steven Phillips

Electrotechnical Laboratory

Abstract

A relation-based theory of cognition proposes that cognitive capacity is limited, in part, by the maximum arity of a relation that can be processed in working memory (Halford, 1993; Halford, Wilson, & Phillips, submitted). Children below age five are limited to binary relations, hence have great difficulty on transitive inference tasks, which require integration of two binary relations into a ternary relation. This theory attempts to integrate cognitive and developmental data on the basic of a single metric - relational arity (number of related arguments). However, the lack of formal analysis into relational information involved in cognitive tasks threatens to undermine its utility. I propose using *Natural language Information Analysis Method* from relational database theory to analyze relational information in cognitive tasks. To demonstrate the utility of this method, I analyze two tasks: (1) *simple oddity*; and (2) *dimension abstracted oddity*. The analysis identifies the peak arity of *simple oddity* as binary and *dimension abstracted oddity* (like transitive inference) as ternary. Therefore, the relational theory predicts that *dimension abstracted oddity* cannot be performed until the median age of five years, while *simple oddity* can be performed earlier. The analysis also suggests variations on these tasks, and the peak arity for each variation is examined. [1]

Introduction

A relation-based theory of cognition posits that cognitive ability is limited, in part, by the maximum arity of relations that can be processed in working memory (Halford, 1993; Halford et al., submitted). For example, children below age five are said to be limited to binary relations. Hence, they have difficulty performing transitive inference (e.g., the inference that a *is-taller-than* c, given that a *is-taller-than* b and b *is-taller-than* c) as it requires integrating two binary relations into a ternary relation. Children below eleven (limited to ternary relations) have difficulty on the balance-scale task (i.e., given a beam and two weights $[w_1, w_2]$ at distances $[d_1, d_2]$ from a pivot point, determine whether the beam will balance, tip left, or tip right), which requires a quaternary relation.

An extensive review of the experimental literature has shown that many tasks attainable by a particular age group can be categorized on the basis of common relational arity Halford (1993), Halford et al. (submitted), Andrews (submitted). Examples of reasoning tasks incorporated by the theory include: class inclusion, which consists of a superordinate class, a subclass and its complement (e.g., fruit, apples and non-apples); various configural

[1] I thank Graeme Halford and Bill Wilson for many important discussions related to this work. I also thank Motoi Suwa for supporting this research, and Kazuhisa Niki, Kazuo Hiraki and other members of the Cognitive Science Section of ETL for their comments. I also wish to express my appreciation to an anonymous reviewer of this paper who suggested its new title and provided many thoughtful comments and criticisms. This work was supported by a Science and Technology Agency (STA) fellowship.

Correspondence can be sent to: Steven Phillips, Information Science Division, Electrotechnical Laboratory, 1-1-4 Umezono, Tsukuba, Ibaraki 305 Japan. Email: stevep@etl.go.jp.

discrimination tasks such as conditional discrimination and negative patterning, which are analogous to *exclusive OR*; and others (Halford, 1993). The theory has also been applied to language comprehension, and planning tasks, such as the *Towers of Hanoi* (Halford et al., submitted).

The essential point of the theory is that relational arity is an important factor in the ability of subjects to perform cognitive tasks. Consequently, tasks that require processing relations of the same arity should impose the same degree of difficulty, other factors being equal. Tasks of known difficulty form a benchmark against which new tasks are compared. For example, given experimental results suggesting transitive inference is difficult for children below age five years, and that transitive inference requires processing a ternary relation, then other tasks requiring ternary relations should also be difficult for children below age five.

The question as to whether young children can perform transitive inferences has been the subject of some dispute (see Halford, 1993, for a review). Recently, apparent demonstrations of transitive inference in 3- and 4-year-olds by Goswami (1995) have challenged the claim that transitive inference cannot be performed before the median age of five. Although, the generality of these results was subsequently questioned by Halford et al. (submitted). Importantly, a deeper experimental investigation of the processes involved in transitive inferences revealed that children below the age of five could perform transitive inference only in the case where the task could be decomposed into two serial steps of one binary relation each (Andrews & Halford, submitted). Performance for 4-year-olds dropped to chance level when the task demanded integration of both binary relations into a single ordered triple (ternary relational instance) in a single step. The results suggest that although young children may have some understanding of the concept of transitive inference (in that such inferences were performed by serialization), they lack the capacity to perform transitive inferences in the general case. It gives strong support for the claim that young children cannot process ternary relations.

The appeal of the relational theory is that it integrates a range of cognitive behaviour on the basis of a single metric - relational order (arity), which is closely related to the number of interacting variables in an experimental design (Sweller, 1993). Yet its weakness is that in the psychological literature, the concepts of relation and relational order have a wide interpretation - one person's quaternary relation may be another person's ternary relation.

In the computer science literature, however, the concepts of a relational data structure, its associated operators, and methods for analysis (e.g., what constitutes a relation of a particular arity) are well-defined and formalized (Codd, 1990; Halpin, 1995). Of course, the goals of the computer scientist differ from the cognitive scientist, but they share a common ground: a relational theory for information storage and process. Thus, a potentially fruitful line of research is to incorporate the formal concepts and methods from relational database theory towards understanding the sources of relational complexity in cognitive tasks. The working hypothesis of this work is that:

> a significant potion of cognitive behaviour (an interaction between an architecture and an environment) is relation-based; together the cognitive architecture and its environment constitute a relational information system, which can be modeled, analyzed and tested using methods from relational database theory.

Halford's theory identifies relational arity as a metric for complexity in cognitive tasks. One technique from relational database theory called *Natural language Information Analysis Method, NIAM* (Halpin, 1995) is used to decompose information systems into an essential collection of connected relations. If one treats cognitive behavior as a relational information system, then cognitive tasks can be analyzed using NIAM into an essential collection of relations. From there one can identify the peak complexity (maximum arity of a relation) that must be processed to perform the task of interest.

To demonstrate the utility of NIAM, transitive inference, simple oddity, and dimension abstracted oddity are analyzed here. The point will be that the analysis method clarifies the source of relational complexity in these tasks. It identifies dimension abstracted oddity as having the same peak complexity as transitive inference - it requires integration of two binary relational instances into a ternary relation instance, where simple oddity does not require integration. Therefore, Halford's theory predicts that simple oddity can be performed before the median age of five years, whereas dimension abstracted oddity cannot.

Relational complexity and working memory capacity

Relational complexity is closely linked to the number of interacting variables. Intuitively, a decision based on one variable requires less "effort" than one based on two variables, which in turn requires less "effort" than one based on three variables, etc. For example, if a decision on rented accommodation was based solely on "cost" (by simply comparing prices), then this would be a relately effortless decision compared to one based on "cost" and "floor space". Since in the second case we would need a way of weighting both "cost" and "floor space" variables. Similarly, if we include further variables such as "distance to work" and "surrounding environment", the decision becomes increasingly effortful, since each new variable interacts with previous variables requiring further consideration.

The essential idea in Halford's theory is that the capacity to perform cognitive tasks is limited, in part, by the number of independent interacting pieces of information that can be processed in working memory (i.e., the maximum relational arity). Higher arity relations can be "chunked" into lower arity relations, thus circumventing a working memory capacity limitation. However, chunking requires a strategy, which must be learned. Therefore, on "novel" tasks where the availability of such strategies is highly unlikely, maximum relational complexity will be the determining factor. It will constitute the peak processing load for a given task.

In understanding Halford's theory, it is important to distinguish two dimensions of potential development in working memory: (1) the number of independent objects; and (2) the number of interacting objects that can be processed. Traditionally, theorists have pointed to the first dimension as the factor that limits subject's ability to perform tasks. Halford (1993), by contrast, identifies the second dimension as another critical factor.

Traditionally, using the computer metaphor, one thinks of memory capacity as the number of slots or positions where chunks of information (e.g., letters, numbers, symbols) are held for processing. A more appropriate metaphor illustrating Halford's theory is to think of working memory as a slotted box, such as the ones used to store compact discs. Not only is there a limit to the number of slots, but there is also a limit to the height of the box corresponding the arity of relational information that can be stored for subsequent processing. A shallow box is sufficient to hold mini-discs (corresponding to instances of unary relations, say), but not sufficient to hold compact discs (corresponding to instances of binary relations). Likewise, a deeper box sufficient for holding compact discs, is not necessarily sufficient for holding laser discs (ternary relational instances).

A cognitive task may involve a series of operations on a number of relations of varying arities. "Peak" relational complexity refers to the maximum arity of the relations computed in a task. Capacity refers to the maximum arity of a relation that can be processed in working memory. Using the "slotted-box" metaphor, capacity will be the depth of the box, not the number of slots. If peak relational complexity exceeds working memory capacity then the task cannot be performed above a chance-level baseline (Andrews & Halford, submitted).

Relational complexity, however, does not necessarily equate to apparent relational arity. Some apparently higher arity relations can be decomposed into a collection of lower arity relations without loss of information. Conversely, some tasks require integration of lower-arity relations into higher-arity relations to extract the target information. As shall be demonstrated here, determination of relational complexity requires analysis of the information relevant to the task.

What constitutes relational complexity?

The attraction of the relational complexity approach of Halford's is that data indicating capacity limitations from many developmental and cognitive psychological tasks have been analyzed and shown to be consistent within this framework (Halford, 1993; Halford et al., submitted). However, there is some argument as to what constitutes the arity of a relation, and what contributes to the complexity of processing a relation (anonymous reviewers of Halford, Wilson, & Phillips, submitted).

For example, in the *oddity discrimination* tasks, it is not clear what part of these tasks constitutes the peak relational complexity. In *simple-oddity discrimination* a subject is required to determine which object differs from the other objects (e.g., given two triangles and a circle, the circle is the *odd* object). In *dimension-abstracted oddity*, objects vary along two or more dimensions (e.g., shape and colour), only one of which determines the *odd* object. For example, given a green circle, a blue triangle and a blue square, the green circle is the *odd* object because this object is green, while the other objects are blue. Shape does determine the odd object because every object has a different shape.

Simple oddity tasks can be performed by children below five, whereas *dimension-abstracted oddity* cannot. Although the complexity of discrimination tasks has not been completely analyzed, Halford (1993) suggested that this difference could be due to *simply oddity* requiring only the single binary relation "Difference(Object1, Object2)", whereas *dimension-abstracted oddity* requires integration of the two binary relations "Same(Object1, Object2)" and "Difference(Object1, Object3)" into a ternary relation. Since children below five are limited to binary relations they can perform the first task, but not the second.

A more complete analysis of relational information must specify the relations and operations needed to determine the *odd* object. When specified we see that determining relational complexity is a complex issue. In *simple oddity*, for instance, the *odd* object is an element of the complex relation: And(Different(Odd, Object2), Same(Object2, Object3)), which is constructed from binary relations "And", "Same" and "Different". Now, what is the relational complexity of this example? If we treat relational complexity as the arity of the relation with the greatest number of arguments then it is binary. Yet, the "And" relation is operating over two binary relations of two arguments each, and so can be considered as a quaternary relation. Yet again, two of these arguments are over the same object (Object2). Removal of the redundant argument results in a ternary relation. So, for one task we have three different measurements of complexity.

Relational complexity is well-defined in the relational database literature as the minimum arity that a relation can be reduced to without loss of information (Codd, 1990). Furthermore, analysis methods such as NIAM (Halpin, 1995) can be used to formally analyze existing information systems into a minimum set of minimal arity relations (called *optimal normal form*) sufficient for representing and processing relational information in some domain. Although NIAM has proven successful in the area of database systems it has not been applied to cognitive systems.

Relational processing and transitive inference

Formally, an n-ary relation is a subset of tuples from the cartesian product of n sets, such that each tuple represents a *truth* about the domain of interest. For example, the binary relation "Taller" may be the set: $R_T(P_1, P_2) = \{(\mathrm{Mark}, \mathrm{Bill}), (\mathrm{Bill}, \mathrm{Tom}), (\mathrm{Bill}, \mathrm{Paul})\}$, where R_T is the symbol of the relation "Taller", and P_1 (P_2) are the first (second) arguments to the relation. If we consider a relation as a table, where rows are instances of the relation and column names are the arguments of the relation (Figure 1(a)), then the relational operator *select* ($\sigma_C R$) returns rows (elements) of the table (relation R) satisfying condition C, and *project* ($\pi_A R$) returns columns with argument name A. Questions such as *who is taller than Bill?* are answered as: $\pi_{P_1}(\sigma_{P_2=\mathrm{Bill}} R_T) \rightarrow$ Mark. That is, by retrieving the row with "Bill" at argument P_2 (Figure 1(b)), and then the element at argument P_1 of that row (Figure 1(c)).

T(P1,P2)	P1	P2
	Mark	Bill
	Bill	Tom
	Bill	Paul

(a)

T(P1, Bill)	P1	P2
	Mark	Bill

(b)

T_{Bill} (P1)	P1
	Mark

(c)

Figure 1: The relation "Taller" depicted as a table (a), and the result of successively applying the operations of select (b) and project (c).

The difficulty with performing transitive inference has been identified with the integration of two binary relations into a ternary relation of ordered triples (Andrews & Halford, submitted). Formally, integration corresponds to performing an *equi-join* operation ($\ominus_{A_1,A_2}$), which joins two relations on the basis of common elements at positions A_1 and A_2 of the first and second relations, respectively. For example, the inference "Taller(Mark, Tom)" is realized as $\sigma_{(P_1=\mathrm{Mark} \wedge P_3=\mathrm{Tom})} R_T \ominus_{P_2,P_1} R_T$. The *equi-join* operator returns relation: $R_{TT}(P_1, P_2 = P_1, P_3) = \{(\mathrm{Mark}, \mathrm{Bill}, \mathrm{Tom}), (\mathrm{Mark}, \mathrm{Bill}, \mathrm{John})\}$, and the *select* operator returns the row with "Mark" ("Tom") in the first (third) position of R_{TT} (Figure 2).

Now, the important point of this formalism with respect to Halford's theory is that it identifies the peak complexity of the transitive inference task as the ternary relation R_{TT} constructed from the binary relation R_T using the *equi-join* operator. If we continue the analysis on the *oddity discrimination* tasks we see that *dimension-abstracted oddity* has the same peak complexity as transitive inference (i.e., requires the same equi-join operation), whereas *simple oddity* does not. Therefore, *dimension-abstracted oddity* should present the same degree of difficulty to children as transitive inference, which is consistent with Halford's (1993) prediction.

At this point, it should be noted that the analysis is not committed to a particular way of implementing, for example, an equi-join operation. Its purpose is to identify common relational data structures and operations. If particular operations have been identified as the source of difficulty in one task (e.g., transitive inference), then they are expected to be the source of difficulty in other tasks requiring the same operations.

NIAM applied to discrimination tasks

The NIAM procedure consists of identifying the relevant atomic components (called entities) and the facts to be stored and processed about these entities. The relevant facts are

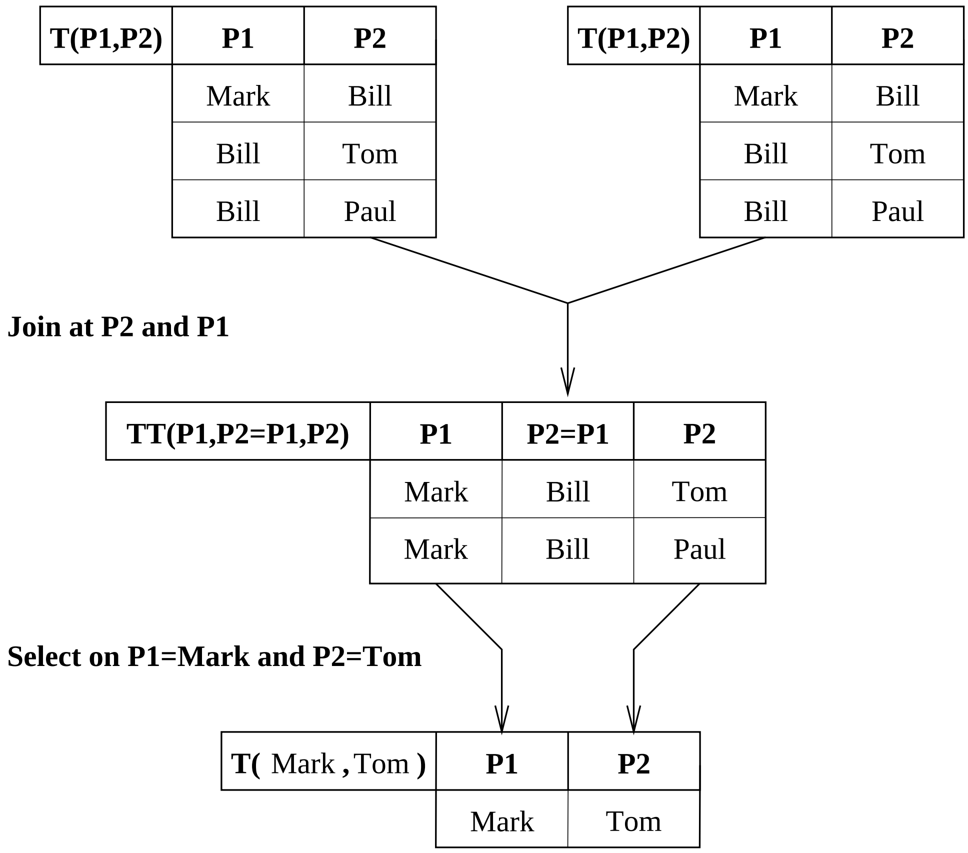

Figure 2: Transitive inference on the "Taller" relation.

identified by writing down statements about the domain of interest and then removing or decomposing those statements so that only:

- facts relevant to the task; and
- facts relating relevant atomic entities are provided.

<u>Simple oddity discrimination</u>

The task requirement of *simple oddity* is to identify the object that differs from the other objects, where the other objects are all the same. For example, given three objects where the first and second objects have a square shape and the third object has a triangle shape, the third object is the odd object as the other two objects are both squares. So, the sorts of entities and relations relevant to this task are object, shape, the number of shapes, and the relations "same" and "different". The first step is to make explicit the facts relevant to this task. These facts are identified in Table 1. (Starred numbers indicate facts from previous facts.)

A number of other facts are possible (Table 2). However, they are not included in the collection of basic facts about this task for several reasons. Firstly, the entity first square is complex in that it consists of the two atomic entities, first object and square. Complex entities must be decomposed into the relevant atomic entities for the task. The relevance of an atomic entity depends on whether it is used in the task. One could, for example, decompose the entity of square into its lines, angles and vertices. However, such a decomposition is unnecessary as these component entities are not relevant to the task. All that is required is for a subject to determine whether two shapes are the same or different. This part of the task itself is potentially difficult. For example, determining whether an object is a 11-pointed

Table 1: Basic facts relevant to the simple oddity task.

No.	Basic facts
1	the first object is a square
2	the second object is a square
3	the third object is a triangle
4	the first object is the same shape as the second object
5	the first object is a different shape to the third object
6	the second object is the same shape as the first object
7	the second object is a different shape to the third object
8	the third object is a different shape to the first object
9	the third object is a different shape to the second object
10*	there are two squares
11*	there is one triangle
12*	the third object is the odd object

star or a 13-pointed star may impose higher cognitive load than say differentiating between a triangle and a square. However, for the purposes of simple oddity discrimination, it is assumed that the materials are such that determining whether two shapes are the same or different imposes little cognitive load.

Table 2: Complex and/or irrelevant facts for the simple oddity task.

No.	Other facts
13	the first square is left of the second square
14	the first object is the same as the second, but is different from the third object

Fact 13 also identifies a spatial relation that is not relevant to the task. Since the odd object does not depend on its position relative to the other two objects, such facts are not included for analysis. Fact 14 is omitted as it can be decomposed, without loss of information, into the two simpler facts: Fact 4 and Fact 5. The point being that by decomposing the task into its simplest and relevant facts, one can identify the minimum relations necessary to identify the target. If one started with complex entities and facts, then there is the possibility of missing a solution that uses less complex relations.

Now, assuming that Facts 1-12 are the only facts relevant to the task, the next step of NIAM is to make explicit the entities and relations. Typically, entities are the nouns and relations are the verbs in each statement. This stage of analysis is potentially the weakest since it relies on common sense knowledge of the relevant concepts[1]; knowledge which may not be available to the subject, particularly in the case of children. However, one of the major strengths of NIAM is that it makes explicit what conditions are being assumed.

The relevant entities are object, for which there are three instances (i.e., first, second and third, which one can simply label as o_1, o_2, o_3, respectively); the shape entity having instances square and triangle; and the number entity (e.g., 1, 2, etc). The relevant relations are: “Same”, “Different”, “Has (shape)”; and “Occurs”. Slightly more formally then the facts are recast into a standard form, such as:

[1]Although, the choices are not completely arbitrary as demonstrated by showing examples of facts not relevant to the task.

1. the object with object-label o_1 is the *Same* as the object with object-label o_2;
2. the object with object-label o_1 *Has* shape with shape-name square;
3. the shape with shape-name square *Occurs* with number of occurrences 3 ; etc.

The conceptual schema that captures these facts is shown in Figure 3(a). Circles indicate entities, which are sets of objects in the domain of interest. Rectangles indicate relations and the number of partitions in each rectangle identifies the arity of the relation. In this case, all relations are binary. The elements in each relation are shown next to the rectangles. The starred box indicates a derivable relation (i.e., a relation constructed from other relations). It is assumed that the counting process does not require more than a rehearsed procedure, which runs through pairs of associations between a number and its successor.

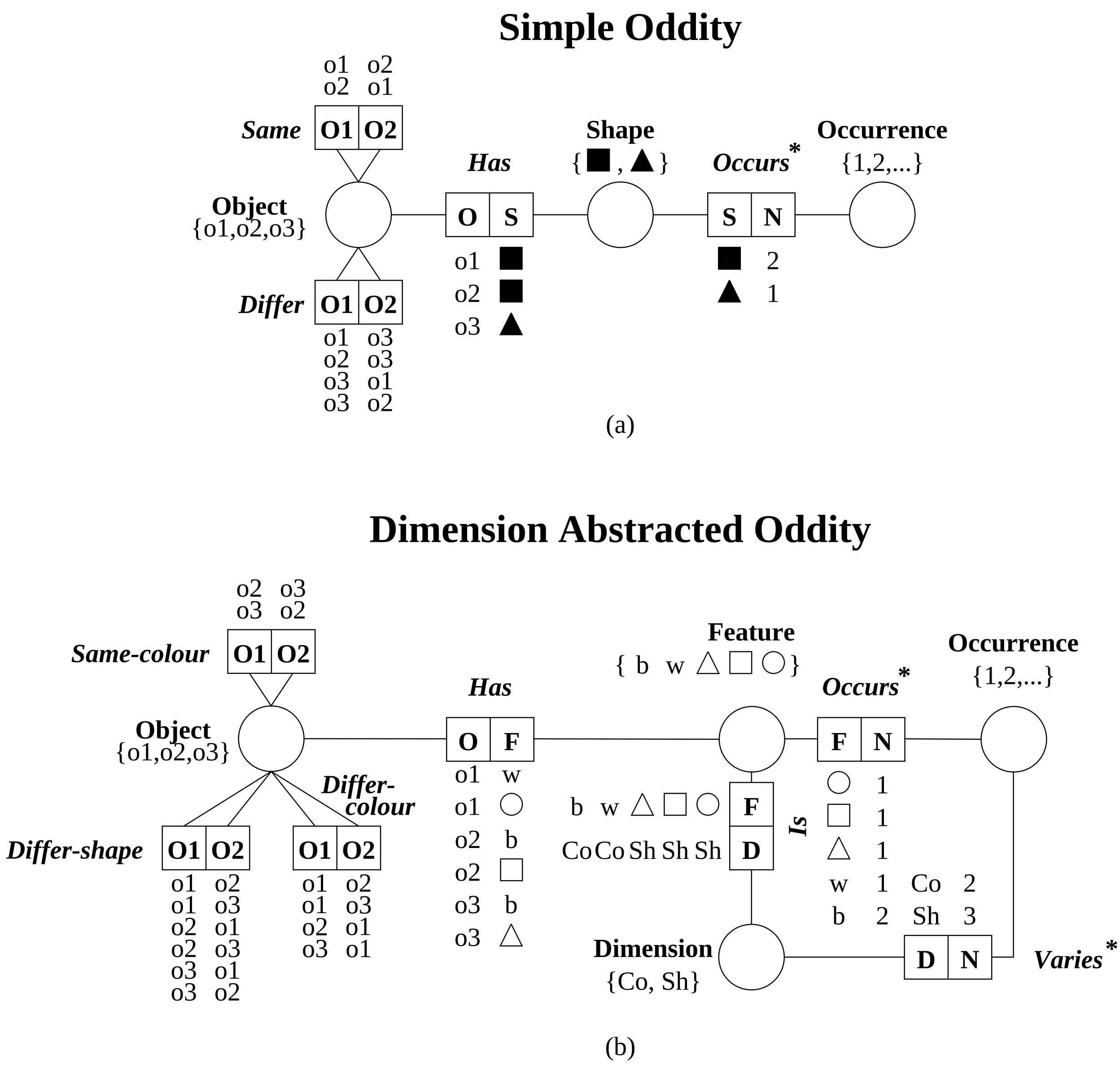

Figure 3: Conceptual schema for *simple oddity* (a) and *dimension-abstracted oddity* (b) tasks.

Previously, I suggested that *simple oddity* could entail a ternary relation. From the relations specified in the conceptual schema one can determine the *odd* object by performing an *equi-join* on the “Different” and “Same” relations resulting in the relation: $R_{DS} = R_D \ominus_{O_2,O_1} R_S = \{(o_3, o_1, o_2), (o_3, o_2, o_1)\}$, and then a *project* on the first argument:

$\pi_{O_1} R_{DS} \to o_3$. (Note that since relations are sets duplicates are eliminated.) However, the conceptual schema also identifies an alternative strategy using the derivable relation "Occurs", which relates shape to the number of times it appears in the task. Using this relation one can avoid constructing the intermediate ternary relation by a two step serial process: 1. $\pi_O(\sigma_{N=1} R_O) \to \triangle$, which returns the shape that occurs once; and 2. $\pi_O(\sigma_{S=\triangle} R_H) \to o_3$, which returns the object that has shape triangle. This two step serial strategy avoids the expensive join operation and requires processing at most a binary relation.

Dimension abstracted oddity discrimination

Dimension-abstracted oddity consists of three or more objects with features that vary along two or more dimensions. The task is to identify the object with a feature that differs from all other objects where those objects share the same feature. For example, given the three objects with feature pairs: (white, circle); (black, square); (black triangle), respectively, the white circle object is odd because its colour is white, whereas all other objects are black. Shape is not the discriminating dimension because it does not partition the objects into two groups where all objects in the same group share the same shape.

Suppose a *dimension-abstracted oddity* task consisting of a white circle, a black square, and a black triangle. The basic facts concerning this task are given in Table 3. The conceptual schema that captures these facts is shown in Figure 3(b). The schema records facts such as: "the Feature with feature-name circle ($\bigcirc$) *Is* a Dimension with dimension-name Shape (Sh)"; "the Feature with feature-name black (b) *Occurs* with number of occurrences 2"; etc. The interesting feature of this schema is that the two step procedure in the previous task cannot be applied since more than one feature has a single occurrence (i.e., $\pi_F(\sigma_{N=1} R_O) \to \{w, \triangle, \square, \bigcirc\}$). The structure of this task requires first determining the discriminating dimension, which is the dimension with only two values (i.e., $\pi_D(\sigma_{N=2} R_I) \to$ Colour), and second integrating the "Is" relation with the "Occurs" relation using the *equi-join* operator (i.e., $\pi_F(\sigma_{(D=\text{Colour} \wedge N=1)} R_C \ominus_{F,F} R_O) \to w$, where w is the *white* feature). The third and final step simply selects the object having the *white* feature (i.e., $\pi_O(\sigma_{F=w} R_H) \to o_1$). The important point to take from this analysis is that *dimension-abstracted oddity*, like transitive inference requires a ternary relation, implying the same level of difficulty for children below age five.

Variations on dimension abstracted oddity

The essential lesson to take from NIAM is that domain complexity depends on the facts that must be operated on. If one alters the facts then task complexity may change. Here, two variations on the dimension abstracted oddity task are examined.

Dimension abstracted oddity with deletion

The analysis of dimension abstracted oddity assumed that the facts remained unchanged throughout the task. One can also consider the effect of task complexity if the subject is permitted to alter (in this case, delete) facts.

Dimension abstracted oddity first requires the subject to identify the discriminating dimension (i.e., the dimension with only two possible features). In the example, this dimension was colour. Colour becomes the cue to the "Is" relation, which returns two possible features: black and white. To decide between which colour, one joins the "Is" relation with the "Occurs" relation along the common position "F" (feature), resulting in the ternary relation $R_{IO}(D, F, N)$. The target is accessed using the additional cue "1" to return instance (colour, white, 1).

Table 3: Basic facts relevant to the dimension abstracted oddity task.

No.	Basic facts
1	the first object is a circle
2	the second object is a square
3	the third object is a triangle
4	the first object is white
5	the second object is black
6	the third object is black
7	a circle is a shape
8	a square is a shape
9	a triangle is a shape
10	white is a colour
11	black is a colour
12*	there is one circle
13*	there is one squares
14*	there is one triangle
15*	there is one white object
16*	there is one black object
17*	there are two colours
18*	there are three shapes
19*	the first object is odd
20	the first object is a different shape from the second object
21	the first object is a different shape from the third object
22	the first object is a different colour from the second object
23	the first object is a different colour from the third object
24	the second object is a different shape from the first object
25	the second object is a different shape from the third object
26	the second object is a different colour from the first object
27	the second object is the same colour as the third object
28	the third object is a different shape from the first object
29	the third object is a different shape from the second object
30	the third object is a different colour from the first object
31	the third object is the same colour as the second object

An alternative strategy, suggested by Graeme Halford[2], is to permit subjects to remove objects in the task not relevant to the solution. In this case, having identified the determining feature as being either black or white, the subject chooses one of them. With the chosen feature, they determine the number of occurrences in the task by selecting from the "Occurs" relation (a single binary relation operation). Now, if the retrieved number of occurrences is not one, then remove all objects with that feature, otherwise the selected feature is the feature that determines the odd object. So, for example, if the subject chooses black, which they determine has 2 occurrences, they remove all black objects. Deletion of these objects leaves only one object which must be the odd object. However, had they chose white, which has only one occurrence, they use this feature to cue the "Has" relation which identifies the first object as the odd object. In both cases, only binary relations are used. There is no joining to form ternary relations. Using a deletion strategy permits the task to be performed with only binary relations. If such a strategy is available to subjects, then Halford's theory predicts that this task can be perform below the median age of five years.

$D > 2$ dimension abstracted oddity

Another variation on this task is the situation where objects vary along three or more dimension. Suppose a task of four objects[3] with size being the additional dimension of variation. Now, the affect of adding a third dimension will depend on the number of values along this dimension that appear in the group of objects to be discriminated.

In the case of the four objects: big white circle, big black square, medium black square, and small black triangle, there are three values along the size dimension (i.e., small, medium and big). The "Varies" relation now contains three pairs: (colour,2),(shape,3),(size,3). However, the addition of the third pair does not change the relational complexity of the task, because the discriminating dimension is still uniquely identified as the dimension with only two values (i.e., $\sigma_{N=2} R_V \rightarrow \text{Co}$). The determining feature is then retrieved by use of the equi-join operator as before. Thus, the relational complexity of this variant is still ternary.

Relational complexity does not preclude other factors contributing to the increased difficulty of a task. One can imagine the task of discrimination being more difficult when objects vary over many dimensions, since there more dimensions must be checked, most of which would be irrelevant. Yet, in this case, the increase in difficulty would be attributable to other factors (e.g., time), which are outside the scope of the relational theory. Simply having to check more dimensions, or more objects takes more time. To take an extreme example, searching for a *needle in a haystack* is no more demanding, cognitively, than looking for a needle in a handful of straws. Yet, people are more likely to fail on the first task either out of lack of patience, or an inappropriate strategy (e.g., *in situ* search).

Another alternative is where the third dimension has only two possible values, as for example in the group of objects: big white circle, big black square, small black square and small black triangle. In this case, the size dimension consists of only two possible values: small and big, yet size is not the discriminating dimension. Although, the relation "Varies" contains the three pairs: (colour,2); (shape,3); (size,2), the discriminating dimension cannot be uniquely identified using the cue "2" (i.e., $\sigma_{N=2} R_V \rightarrow \{\text{Co}, \text{Si}\}$). To identify the determining feature, one must perform an equi-join on relations "Varies" and "Is" at the common position "D" (i.e., $R_V \ominus_{D,D} R_I \rightarrow R_{VI}(N, D, F)$), and then another equi-join on the resulting ternary relation with the relation "Occurs" to construct a quaternary relation

[2] Personal communication

[3] The number of objects does not necessarily change the relational complexity of a task. For example, even if one adds an extra square object to the simple oddity task example, the relational complexity is still only binary.

(i.e., $R_{VI} \ominus_{F,F} R_O \rightarrow R_{VIO}(N, D, F, N'))$. This operation is depicted graphically in Figure 4. Cuing this quaternary relation with the number of variations ($N = 2$) and the number of occurrences ($N' = 1$), returns the instance (2,Colour,white,1), where white is the feature that determines the odd object. Thus, it is expected that children at a median age of five years who can perform $D = 2$ dimension abstracted oddity, cannot perform this variation.

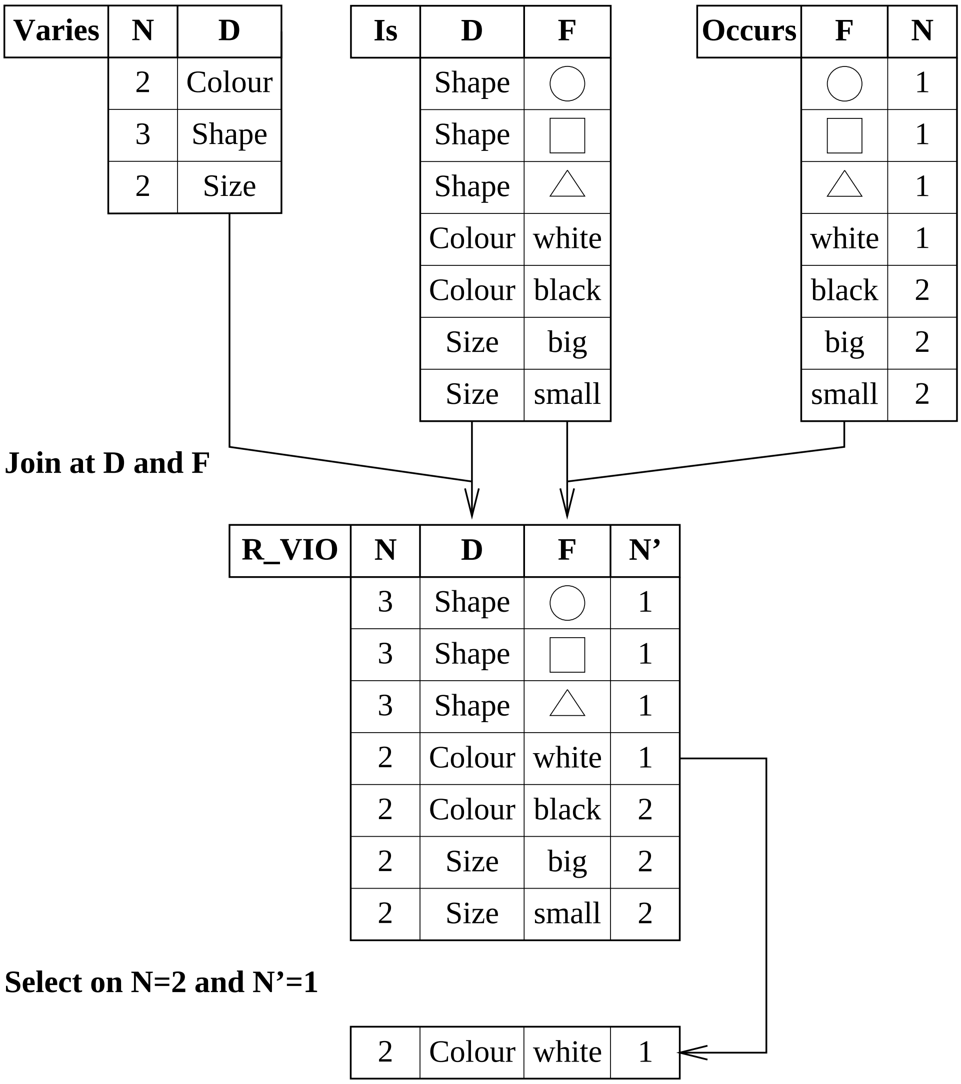

Figure 4: Quaternary relation.

Discussion and conclusion

The purpose of NIAM is to analyze information systems into a minimum collection of minimal arity relations. Once this collection has been identified, one can determine the relational operations needed to extract the target information. Thus, NIAM and the conceptual schemata that are produced by this analysis technique are statements at the algorithmic (relational) level. They are not statements about how relational data structures and processes should be implemented. So, the concepts introduced here to analyze relational complexity in cognitive tasks, should not be confused with the concepts used in relational database technology, which are concerned with specific implementations of relational information systems. Whereas in database design it is appropriate to consider such concepts as balance-trees and hash tables to maximize speed and efficiency, in cognitive modelling and in particular connectionist modelling it is more appropriate to consider concepts such as distributivity and connectivity for neural plausibility (Halford et al., submitted).

The value of this approach has been to identify a common source of relational complexity between transitive inference and dimension abstract oddity tasks, which empirical evidence suggests is the source of difficulty for children below age five (Andrews & Halford, submitted). The work of Andrews and Halford identifies integration of two binary relations into a ternary relation as the source of difficulty for young children on transitive inference. This integration corresponds to performing an equi-join operation at the relational level. Therefore, using ternary inference as the empirical benchmark, other tasks requiring the same operation should yield the same degree of difficulty, other factors being constant. Using NIAM, dimension abstracted oddity was identified as one such task.

One should bear in mind that the conceptual schemata capture information contained and processed in a relational information system. With respect to cognitive behaviour that system constitutes the environment (task materials) and the subject. So, not all the information will reside in the subject's working memory at any one time. Some of this information may simply be present in the environment as task materials, which can be retrieved on demand by appropriately focusing on the relevant materials. So, for example, a simple oddity discrimination task may consist of ten objects (say, nine squares and a circle). This task instance would consist of the binary relation "Has" containing ten pairs. The subject, though, does not have to store every pair in memory. Probably only two pairs need be stored. That is (square, 9) and (circle, 1), from which the subject identifies the determining feature as the circle. The subject then only has to scan the scene for a circle. Each time the subject encounters an object, they compute the relational instance $R_H(o_i, s_j)$. If the current instance contains the shape circle, it is used to determine the odd object, otherwise it can be discarded and the remaining objects considered.

Although the applications of NIAM by the systems analyst have differed considerably from the cognitive scientist, it is entirely appropriate to use this method because they share a common goal: the reduction of information systems into the smallest possible collection of relations. Fulfilling this goal from the analyst's point of view helps enforce data consistency and integrity. From the cognitive scientist's viewpoint, and in particular with respect to Halford's theory, it identifies lower bounds on the maximum arity of relational information inherent in the task. Therefore, if these bounds exceed the capacity of subject's working memory, one can predict that such tasks cannot be performed without some strategy for reducing the maximum arity of relational information. In the previous section one such strategy was discussed. In the dimension abstracted oddity task using deletion of objects, "peak" complexity was reduced from ternary to binary.

Conversely, the analysis also identifies strategies that process relations of higher arity. For example, simple oddity required a ternary relation when the strategy for determining the odd object consisted of integrating the "Same" and "Difference" relations. So, although young children may be at the developmental level that permits them to perform simple oddity, their failures may be due to a strategy that entails processing relations beyond their level. Furthermore, since NIAM is a fact-oriented analysis method, when additional facts are added to the tasks, such as in $D > 2$ dimension abstracted oddity, one can determine the relational complexity of the modified task and make predictions as to subjects' performance.

The weakest point of this analysis method is the first step of identifying the atomic facts. Atomicity is relative to the cognitive system. An individual face contains many relations relevant for discrimination from other faces, yet face recognition is one of the earliest acquired and least cognitively demanding capacities. Even for object categorization, what adults would regard as two separate dimensions of similarity (e.g., colour and size), young children below the age of five integrate into a single dimension from which they determine relative "sameness" (Smith, 1989). In the case of Smith's (1989) results (Experiment 1), NIAM could account for the difference by identifying a binary relation between object pair and an overall-similarity metric for the young children (e.g., Similarity(Pair, Distance)), and a ternary relation between object pair, dimension and a dimension-specific-similarity metric for the older children and adults (e.g., Similarity(Pair, Dimension, Distance)). But, this would beg the question of the psychological reality of the Distance attributes in the two relations.

The purpose of NIAM is not to provide an explanation for the atomic objects of cognitive processes. Its purpose is to provide explicit and formal methods for determining relational information given those atomic objects. Thus, when calibrated against tasks of known difficulty (e.g., transitive inference) it can be used to identify common sources of relational complexity. Relative to the relational theory, then, tasks sharing common relational complexity should present the same degree of difficulty.

Relational systems have been studied extensively within the relational database community. As the analysis in this paper demonstrates much understanding of the relational complexity inherent in cognitive tasks can be gained by treating the cognitive system (architecture plus environment) as a relational system, thereby making the cognitive system subjectable to the formal methods of relational information systems.

References

Andrews, G. (submitted). *Relational complexity as a capacity construct in cognitive development.* Ph.D. thesis, The University of Queensland.

Andrews, G., & Halford, G. S. (submitted). Children's ability to make transitive inferences: The importance of premise integration and structural complexity.. http://www.psy.uq.edu.au/Department/Staff/gsh.

Codd, E. F. (1990). *The Relational Model for Database Management: Version 2.* Addison-Wesley.

Goswami, U. (1995). Transitive relational mappings in 3- and 4-year olds: The analogy of Goldilocks and the three bears. *Child Development*, *66*, 877–892.

Halford, G. S. (1993). *Children's understanding: The development of mental models.* Hillsdale, NJ: Lawrence Erlbaum.

Halford, G. S., Wilson, W. H., & Phillips, S. (submitted). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology.. http://www.psy.uq.edu.au/Department/Staff/gsh.

Halpin, T. A. (1995). *Conceptual schema and relational database design* (2nd edition). Sydney: Prentice Hall Australia.

Smith, L. B. (1989). A model of perceptual classification in children and adults. *Psychological Review*, *96*(1), 125–144.

Sweller, J. (1993). Some cognitive processes and their consequences for the organisation and presentation of information. *Australian Journal of Psychology*, *45*(1), 1–8.