GRANULARITY

Jerry R. Hobbs
Artificial Intelligence Center
SRI International
and
Center for the Study of Language and Information
Stanford University

Abstract

This paper presents a framework for a theory of granularity, which is seen as a means of constructing simple theories out of more complex ones. A transitive indistinguishability relation can be defined by means of a set of relevant predicates, allowing simplification of a theory of complex phenomena into computationally tractable local theories, or granularities. Nontransitive indistinguishability relations can be characterized in terms of relevant partial predicates, and idealization allows simplification into tractable local theories. Various local theories must be linked with each other by means of articulation axioms, to allow shifts of perspective. Such a treatment of granularity must be built into the very foundation of the reasoning processes of intelligent agents in a complex world.

Abstraction

We look at the world under various grain sizes and abstract from it only those things that serve our present interests. Thus, when we are planning a trip, it is sufficient to think of a road as a one-dimensional curve. When we are crossing a road, we must think of it as a surface, and when we are digging up the pavement, it becomes a volume for us. When we are driving down a road, we alternate among these granularities, sometimes conscious of our progress along the one-dimensional curve, sometimes making adjustments in our position on the surface, sometimes slowing down for bumps or potholes in the volume.

Our ability to conceptualize the world at different granularities and to switch among these granularities is fundamental to our intelligence and flexibility. It enables us to map the complexities of the world around us into simple theories that are computationally tractable to reason in. If we are to have a machine of even moderate intelligence, it must have a theory of granularity woven into the very foundation of its reasoning processes. The purpose of this paper is to propose the outlines of such a theory.

Let us assume the world, in whatever complexity the machine is capable of dealing with, is represented in a *global theory*, which we may take to be a first-order logical theory $\mathbf{T_0}$. Our approach to granularity will be to extract from $\mathbf{T_0}$ smaller, more computationally tractable, local theories. Let $\mathbf{P_0}$ be the set of predicates of $\mathbf{T_0}$, and $\mathbf{S_0}$ its domain of interpretation. Suppose a subset $\mathbf{R}$ of $\mathbf{P_0}$ has been determined to be the predicates *relevant* to the situation at hand. We can then define an *indistinguishability relation* $\sim$ on $\mathbf{S_0}$ by means of the following second-order axiom:

(1) $(\forall x, y) x \sim y \equiv (\forall p \in \mathbf{R})(p(x) \equiv p(y))$

That is, x and y are indistinguishable if no relevant predicate distinguishes between them.

In general, of course, it is a very hard problem to determine just which predicates are relevant. We would expect that in the course of planning to achieve a goal or reasoning about a situation, the set of relevant predicates becomes more constrained, and as it does, more entities become indistinguishable "for all practical purposes". In planning a trip from Menlo Park to Los Angeles, for example, the only relevant spatial predicates involve distance from Los Angeles. Two points in the roadway which differ only in their coordinates across and into the roadway, not along it, do not differ in their distance from Los Angeles, and hence are indistinguishable.

$$< x, y_1, z_1 > \sim < x, y_2, z_2 >$$

This indistinguishability relation allows us to simplify the volumetric roadway into a one-dimensional curve.

Simplification

The indistinguishability relation allows us to define a mapping κ that will collapse the complex theory $\mathbf{T_0}$ into a simpler, more "coarse-grained" theory $\mathbf{T_1}$. Let $\mathbf{S_1}$ be the set of equivalence classes of $\mathbf{S_0}$ with respect to $\sim$, and let κ be the mapping that takes any element of $\mathbf{S_0}$ into its equivalence class in $\mathbf{S_1}$. That is,

Originally appeared in *Proceedings of IJCAI-85*, 432-435.

$$(\forall x, y) x \sim y \equiv \kappa(x) = \kappa(y)$$

Similarly, we can define the mapping κ on the relevant predicates; $\kappa(p)(\kappa(x))$ is true if and only if $p(x)$ is true. Because of (1), this is well-defined. Then let $\mathbf{P_1}$ be the set of predicates $\{\kappa(p) | p \in \mathbf{R}\}$.

Consider an example of collapsing a complex theory into a more computationally tractable theory. Suppose we have a full-blooded theory of the world within which there are a large number of agents and objects, times are real numbers, and places are points in Euclidean 3-space. There are a number of events and predicates to describe the events. Thus, to say that there is a moving event e by x of y from z to w that occurs during the interval T, we can write

$$move(e, x, y, z, w) \wedge duration(e, T)$$

Suppose we are only concerned with discrete, nonoverlapping events in the blocks world, and wish to model it with the situation calculus [4]. We may define a coarse-grained theory $\mathbf{T_1}$ which has in its domain a single agent A. The objects will be only the table and blocks. The relevant places are squares on a 100x100 board, whose lower left corner we will assume to be at the origin. In addition, there is an entity we will call EE, for "everything else". The mapping κ may then be defined as follows:

$\kappa(A) = A$, for the agent A.
$\kappa(x) = EE$, for all other agents x.
$\kappa(TAB) = TAB$, for the table TAB.
$\kappa(b) = b$, for all blocks b on the table TAB.
$\kappa(x) = EE$, for all other objects x.
$\kappa(< x, y, z >) = < floor(x), floor(y) >$, for $0 \leq x, y \leq 100, z = 0$.
$\kappa(l) = EE$, for all other locations.
$\kappa(move) = move$.

We can map continuous time into the relevant discrete times by mapping every instant of time into the end point of the time interval associated with the last event that A was the agent in and mapping every instant of time before the first such event into the starting point of that event. Where events are viewed in theory $\mathbf{T_0}$ as taking place across a time interval, they will be viewed in theory $\mathbf{T_1}$ as happening instantaneously at the end of that interval. Thus,

$$\kappa(\lambda e, T[duration(e, T)]) = \lambda e, T[at\text{-}time(e, end(T))]$$

For every time t in theory $\mathbf{T_1}$, block B_1 will be on block B_2 if and only if B_1 is on B_2 at time t in theory $\mathbf{T_0}$. But we could not have specified that the location of blocks at other times in $\mathbf{T_0}$ was relevant, for that is not preserved by κ, for example, during an interval when a move is being made. By this process we have abstracted from a complex world the classical blocks microworld of AI.

There is evidence from linguistics that our ability to simplify our view of the world in this way is crucial in our use of language. Nunberg [5] has argued that an account of definite reference and the use of the phrase "the same" is greatly simplified if we assume that people construct "local models" of what is going on and treat definite reference as reference to items in the local model. Thus, we can say to a friend,

> Don't buy that car. I used to own the same car and it gave me nothing but trouble.

when we mean not that we once owned the same car token but the same car type. In this situation the only relevant distinguishing predicates concern such things as make, model, and year, and this imposes an indistinguishability relation that does not distinguish among certain car tokens. However, we cannot say

> A 1965 Mustang jumped the center divider of Highway 101 and hit the same car.

to mean that it hit another 1965 Mustang. Among the relevant predicates in this example are speed and direction of travel and they distinguish between the car tokens.

Idealization

Frequently, an indistinguishability relation is given *a priori*. It is apparently true, for example, that people are able to distinguish between temperatures that are two degrees Fahrenheit apart, but not one degree Fahrenheit. We can distinguish between 58° and 60°, but not between 58° and 59°. Visual acuity and color discrimination provide other examples. This is not a shortcoming in our capabilities, but rather one of the ways we are attuned to the aspects of our environment that are most likely to be relevant to our interests.

Typically, this kind of indistinguishability relation on a set $\mathbf{S}$ is formalized by assuming there is a mapping f from $\mathbf{S}$ to the reals and defining $\sim$ as follows:

(2) $(\forall x, y) x \sim y \equiv |f(x) - f(y)| < \epsilon$

for some ϵ [6]. In the case of temperatures, the Fahrenheit scale maps temperatures into the reals and ϵ is somewhere between 1 and 2.

The first problem these observations raise for our framework is how to relate this notion of indistinguishability to some set of relevant predicates. One way is to say that *partial* predicates may also be relevant. Then we may revise the general definition (1) of indistinguishability to be the following:

(3) $x \sim y$ if and only if for all $p \in \mathbf{R}$, if $p(x)$ and $p(y)$ are both defined, then $p(x) \equiv p(y)$.

That is, to distinguish between x and y, we must find a relevant partial predicate p that is true for one and false for the other. Definition (2) is a special case of definition

(3), where ε is the size of the gaps between the values for which the relevant partial predicates are true and those for which they are false. In the temperature example, let the relevant partial predicates be "x is around $t°$" for all real numbers t, which are true for, say, $t - 3 < x < t + 3$, false for $x < t - 3 - \varepsilon, x > t + 3 + \varepsilon$, and undefined otherwise, where ε is between 1 and 2. Then any two temperatures that are only 1° apart cannot be distinguished by any of the relevant partial predicates. Briefly, if we represent the fuzzy judgments people make by means of partial predicates, then ε in definition (2) can be seen as a measure of their undefinedness.

The next problem is that the indistinguishability relation defined in (1) is transitive, whereas that defined by partial predicates as in (3) need not be. Thus we cannot use indistinguishability to collapse our global theory into a local theory, for the collapsing mapping κ as given above is not well-defined. How is simplification to be achieved with this new definition of indistinguishability?

Suppose we are given a set $\mathbf{R}$ of relevant partial predicates. Consider the set $\mathbf{S}_0'$ of unambiguous elements of $\mathbf{S}_0$, that is, the elements for which every predicate in $\mathbf{R}$ is defined. (This subset may be empty, in which case all the burden falls on the second step.) For $\mathbf{S}_0'$, the indistinguishability relation $\sim$ is transitive, and hence we may take $\mathbf{S}_1$ to be the set of equivalence classes imposed on $\mathbf{S}_0'$ by $\sim$, and define κ as the function that maps every element of $\mathbf{S}_0'$ into its equivalance class in $\mathbf{S}_1$.

Suppose, for example, we are interested only in a rough characterization of temperatures as being in the 50's, or in the 60's, etc. We have here a set of predicates that are not overlapping, i.e., if $p(x)$ then $\neg q(x)$ for every other relevant predicate q in $\mathbf{R}$. Then for those temperatures t that are unambiguously in the 50's, that is, the temperatures from 51 to 58, we can define $\kappa(t)$ to be "50s".

We should extend κ to the ambiguous elements $\mathbf{S}_0 - \mathbf{S}_0'$, so that all circumstances will be covered by the theory. It seems natural for $\mathbf{S}_1$ to inherit indistinguishability from $\mathbf{S}_0$. Since 59° and 60° are indistinguishable in $\mathbf{T}_0$, the 50's and the 60's are indistinguishable in $\mathbf{T}_1$. The residual fuzz never disappears. If we further collapse $\mathbf{S}_1$ into a set $\mathbf{S}_2$ consisting of the 00's, the 100's, etc., then the 00's will be indistinguishable from the 100's since 99° is indistinguishable from 100°. However, we realize that inherited indistinguishability becomes less and less appropriate as we collapse to simpler theories. To eliminate the indistinguishability, we simply stipulate that all elements of $\mathbf{S}_1$ are distinguishable. This is the process of *idealization*.

We thus define κ arbitrarily on the ambiguous elements, respecting, however, the relevant structure of $\mathbf{T}_0$, e.g., the order, which can be defined on $\mathbf{S}_1$ in the obvious way. Thus we would impose the condition that

$$\neg(\exists x, y) x < y \wedge \kappa(x) > \kappa(y))$$

We might map 59 into 50s and 60 into 60s, even though they are indistinguishable in $\mathbf{T}_0$, and declare by fiat that the 50's and the 60's are distinguishable. We thereby sacrifice the tight connection between our local theory and the overarching global theory, as one always does in idealization, but it may be that the resulting local theory is clean enough to make this sacrifice worthwhile.

The idealization should be *faithful* to the global theory, insofar as possible. One measure of the faithfulness of an idealization is the proportion of the entire set $\mathbf{S}_0$ that it was necessary to be arbitrary on, elements x where

$$(\exists y) x \sim y \wedge \kappa(x) \neq \kappa(y)$$

In the temperature example $\mathbf{S}_1$ we have been arbitrary on 20% of the set, so that the idealization is moderately faithful. For $\mathbf{S}_2$, the ambiguous areas are only 2% of the entire set, so the idealization is quite a bit more faithful. The aim in defining κ is to construct a useful theory while at the same time maximizing the faithfulness of the idealization, by this or more sophisticated measures.

This approach to the fuzzy quality of granularity, using partial predicates and idealization, contrasts with a treatment that makes the truth of the relevant predicates a matter of degree. Our approach has the disadvantage of forcing nonintuitive sharp distinctions between the areas where a predicate is true, false and undefined. On the other hand, it gives us a discrete, propositional system that is often finite and even small, and that allows both subtlety of expression and tractable computation.

Articulation

People not only view the world at different granularities. They translate easily among the granularities as needs dictate. Therefore, a theory of granularity must say something about how various local theories *articulate*[1] with each other. There has been a certain amount of work in AI on this problem – research on hierarchical problem-solving in expert systems [1] and on hierarchical planning [7]. When we move from one level of a hierarchy to the level below, we are moving from a coarse-grained local theory to a more fine-grained local theory, and the axioms that specify the decomposition of coarse-grained predicates into fine-grained ones constitute the articulation between the two theories. The articulation can often be quite complex. For example, at the granularity appropriate for the commander of a ship, an event might be thought of as an increase on a continuous speed scale. For the officer in charge of the engine room, the same event might have to be conceptualized in terms of the number of boilers that must be operating.[2] There are some general things one can say

[1]Etymology: from the Latin for "joint", from the Indo-European for "to fit together".

[2]I am indebted to Bruce Roberts for this example.

about articulation, but it is largely a matter of spelling out the particular cases in the knowledge base.

Much of our knowledge is grain-dependent. In the knowledge bases we build as we axiomatize commonsense knowledge [3], [2], grain-size must be an explicit argument of many predications. It is first of all required when we are stating axioms that relate how a phenomenon is seen at two different granularities. We have already seen one example of such a relation – the relation between *duration* in continuous time and *at-time* in discrete time in the situation calculus example. Another example comes from the naive physics of materials. What appears as the bending of a flexible object at one granularity can be viewed at a finer granularity as a stretching and a compressing.

In addition, grain-size must sometimes be mentioned explicitly to prevent us from falling out of the region where the local theory applies. For example, the notion of "substance" is a grain-dependent one. Water is water down to the molecular level, but sand is sand only down to the size of a grain of sand, the grain size associated with succotash is somewhat larger than an individual lima bean, and the grain size of traffic is larger than an individual car. The granularity must be explicitly represented for substances in order to avoid paradoxes of infinite divisibility. We can express as follows the axiom that states that if a substance p has an indistinguishability relation $\sim$ determined by the characteristic granularity of p, then a piece of p has proper parts which are of the same substance, provided it has two distinguishable points:

$$(\forall x)(p(x) \wedge (\exists y_1, y_2)(in(y_1, x) \wedge in(y_2, x) \wedge y_1 \not\sim y_2) \\ \supset (\exists z)(partof(z, x) \wedge z \neq x \wedge p(z)))$$

The qualification saves us from paradox.

Finally it should be pointed out that the whole issue in the philosophy of science of the reducibility of one scientific theory to another is an issue of articulation.

Intelligence

This approach to granularity suggests an intriguing view of the intelligence that people have and that intelligent machines will have to have. It is that our knowledge consists of a global theory together with a large number of relatively simple, idealized, grain-dependent, local theories, interrelated by articulation axioms. In a complex situation, we abstract the crucial features from the environment, determining a granularity, and select the corresponding local theory. This is the only computation done in the global theory. The local theory is then applied in the bulk of the problem-solving process. When shifts in perspective are required, when we must translate the problem from one local theory to another, articulation axioms are used.

One consequence of this view is that work done in "neat" domains, even on toy problems in microworlds, is not necessarily wasted when our ultimate concern is the real world in all its complexity. It may be that this work results in the discovery of local theories that will be essential elements in the overall reasoning process.[3]

One cause of despair in artificial intelligence has always been the lurking fear that when we scale up to knowledge bases of the size that is clearly required for intelligent behavior in a complex world, all of the methods we have been developing will become computationally intractable. Considerations of the uses to which the notion of granularity can be put in knowledge bases and reasoning processes suggest that this need not be the case.

Acknowledgments

I am indebted to Johan van Bentham, Henry Kautz, Pat Hayes, Marilyn Stelzner, Sandy Pentland, Marc Dymetman, and Mike Georgeff for valuable discussions during the course of this research. The research was supported by NIH Grant LM03611 from the National Library of Medicine, by Grant IST-8209346 from the National Science Foundation, and by a gift from the System Development Foundation.

References

[1] Davis, Randall, 1982. "Expert Systems: Where are We? And Where Do We Go from Here?" , *The AI Journal*, vol. 3, no. 2, pp. 3-22.

[2] Hobbs, Jerry R., ed., 1985. "Commonsense Summer: Final Report", SRI International.

[3] Hobbs, Jerry R., and Robert C. Moore, eds. 1985. *Formal Theories of the Commonsense World*, Ablex Publishing Company, Norwood, New Jersey.

[4] McCarthy, John and Patrick Hayes, 1969. "Some Philosophical Problems from the Standpoint of Artificial Intelligence", in Donald Michie and Bernard Meltzer, eds.,*Machine Intelligence 4*, Edinburgh University Press, Edinburgh, Scotland.

[5] Nunberg, Geoff, 1985. "Individuation in Context", *Proceedings, West Coast Conference on Formal Linguistics*, Santa Cruz, California, pp. 203-217.

[6] Roberts, Fred S., 1973. "Tolerance Geometry", *Notre Dame Journal of Formal Logic*, vol. 14, no. 1, pp. 68-76.

[7] Sacerdoti, Earl, 1977. *A Structure for Plans and Behavior*, Elsevir North-Holland, Inc., New York.

[3] A form of this view was prevalent at MIT in the early 1970's.