CHAPTER NINE

TOWARDS SPECIALIZATION

A goal of our theory is to understand how "higher" levels of perception might emerge from "lower" ones, i.e., to understand "perceptual hierarchy." In this chapter we discuss this notion and describe a possible model of it called *framework specialization*. We illustrate framework specialization with two examples: the incremental rigidity scheme of Ullman (1983) and "specialized chain bundles." Our presentation is neither complete nor rigorous, but is an extended speculation guided by work in progress.

1. Introduction to specialization

Our approach to the study of perceptual hierarchy is illustrated by a question: Under what conditions does an ensemble of participators in a fixed reflexive framework Θ give rise to a "higher" level observer or class of observers? We believe it is misguided to restrict attention to answers which implicitly postulate a deterministic or reductionistic relationship between the ensemble and the new observers, e.g., to answers postulating that the new observers are unions or products of the participators in the ensemble. Instead, we seek an answer which exploits the fundamental character of observers: observers perform inferences which are not, in general, logically determined by the premises. In our search for a nonreductionistic answer, we have been guided by four key ideas.

Idea 1. The premises of the new observer should be deducible in some manner from the conclusions of the participators in the ensemble.

The appeal of this idea is that it connects the ensemble and the new observer nontrivially, but it also connects them nonreductionistically. For in this case the ensemble determines only the premises of the observer it gives rise to, not its conclusions. Many different observers can be constructed having the same space of premises Y.

The ensemble which gives rise to a new single observer in this way we call an ***instantiation*** of that observer. We do not call it ***the*** instantiation, for it is likely that a given observer can have many different instantiations. The resulting observer we call a ***specialization*** of the ensemble. Again, we do not call it ***the*** specialization, for a given ensemble is likely to have many different specializations. More generally, we will say that a class of inferences A is an ***ascendant*** of a class of inferences B if all premises for inferences in A are deductive consequences of the conclusions of the inferences in B.

Idea 2. If the premises of the specialized observer arise from the conclusions of the instantiating ensemble, then these conclusions should be reliable.

If we want to build higher levels of perception from lower levels then we want the lower levels to be secure before we start building. In chapter eight we discuss precise conditions in which the perceptual conclusions of a dynamical ensemble of participators are matched to the reality observed. The strongest such conditions we call "stably true perception" and "stably true perception in the limit" (**8**–6.8). For the conclusions of the participators to be reliable in these strong senses the dynamics in which they participate must have a stationary measure ν. The conclusions of the participators are then derived from this stationary measure ν via the rcpd construction $m_\pi^{\mathcal{D}\nu}$ (**8**–6.4, **8**–6.6). The stationary measure can be viewed as describing stabilities of the asymptotic behavior of the participator dynamics. Thus the conclusions of the participators, to be reliable in a strong sense, are derived from stabilities of the asymptotic behavior of their dynamics.

In keeping with Idea 1, a channeling to a specialized observer should occur as a result of channelings in Θ to the participators of its instantiating ensemble. Furthermore, to maintain consistency with our nondualistic semantics, the objects of perception of a specialized observer should be other specialized observers. This leads to:

Idea 3. The channeling between two specialized observers is expressed by an interaction between the observers' instantiating participator ensembles, assuming these ensembles to be in the same framework Θ.

One implication of this is that a single channeling, i.e., a single instant of time, at a specialized level may involve infinitely many channelings at the instantiating level, i.e., at the level of Θ. Such an interaction perturbs the asymptotic behavior of each instantiating system. The asymptotic behavior of each system in isolation must be stable in order to make any sense of the perturbation. Granting this stability, if the perturbed asymptotics has sufficient regularity, each system can encode information about the other system which caused the perturbation.

This brings us to the fourth main idea:

Idea 4. The premise of a specialized observer's inference is a stable perturbation of the asymptotics of the observer's instantiation, a perturbation which results from an interaction with another participator system.

Up to this point we have not given a formal definition of "perceptual hierarchy", i.e., of what it means for one inferencing system to be at a higher level than another. One notion of hierarchy would be a set together with a partial order on it. However, there is no reason to suppose that the intuitive idea of specialization outlined above can be so expressed. If A is a specialization of B, and B is a specialization of C, should one suppose that A *must* specialize C? Is it possible that a chain of such specializations might eventually fold back to its origin? Should one replace a partial order with a more local notion of order?

It is clear that we need a more precise understanding of the information flow from a given ensemble's conclusions to its specialization, as mentioned in Idea 1 above. For example, following upon the discussion after Idea 2, it may be possible to deduce the stationary measure of the instantiated ensemble from the set of the latter's conclusion measures, given that some or all of the participators enjoy true perception of their ambient dynamics. In any case, in order to formally develop the above ideas we will assume this, so that Idea 1 may be expressed in the following form:

Idea 1 bis. The premises of the specialized observer should be deducible in some manner from the stabilities of the dynamics of the participators in the instantiating ensemble.

In what follows we reintroduce these ideas in a more formal setting.

2. Hierarchical analytic strategies revisited

We now consider how to construct a formal model of a perceptual hierarchy. In **4**–5 we suggest that the hierarchy arises from an analytic strategy which decomposes the interactions of complex systems into strata, or levels, and which describes the passage of information between strata. Within each level the interaction appears to be homogeneous, i.e., to involve like entities. Within a given system, and at a given level, we say that the relevant entities together constitute the "representation" of that system at that level. Similarly, the total interaction of two complex systems "expresses" itself at a given level by means of the interaction of each system's representation at that level. But there is more to the total interaction than this: within each system information flows between strata. This flow determines the hierarchical relationship on the collection of strata.

When we introduced the notion of hierarchical analytic strategy in **4**–5 we spoke of "entities of like nature" which are irreducible or indecomposable at a given level. The fundamental hierarchical relation holds between this level and another "lower" level at which each entity has its own representation, a representation which provide a first order decomposition of the entity. The hierarchical connection between these two levels is expressed by a canonical form for the passage of information from the constituents of the lower level representation of the entity, to the entity itself at the higher level. And the information which propagates in this canonical way arises from the interactions between these constituents.

This is where, in a participator-dynamical model, the ideas of specialization fit in. In the model we develop here, reflexive observer frameworks represent the possible hierarchical levels, and participators on a framework are the "irreducible entities" at that level.

Specification 2.1. To give a model of a perceptual hierarchy based on participator dynamics is to specify the following:

(i) a form in which complex systems are represented by participators in given frameworks, i.e., at given levels of the hierarchy;

(ii) a form in which interaction of several such systems is expressed at a given level;

(iii) a canonical form for the passing of information between hierarchically related levels in a single system;

(iv) a manner in which an interaction (as described in (ii)) among several systems generates information which then propagates (within each of the separate systems) via the connection described in (iii).

Here is a more detailed proposal for such a model. First consider (i) of 2.1. We let the expression of a complex system **S** at the level of the hierarchy corresponding to the reflexive framework $\Theta = (X, Y, E, S, \pi_\bullet)$ be an ensemble of participators on Θ together with a τ-distribution on Θ, satisfying a ***permissibility condition*** (discussed below). We call these (participator, τ)-data the ***level*** Θ ***expression of*** **S**. It follows that to a complex system, when considered in isolation, are associated participator dynamics in various frameworks. These participator dynamics are the expressions of the system at the various levels of the hierarchy. Suppose that the level Θ expression of **S** is the participator ensemble $(\xi_i, \{Q_i(n)\}_n, \{\eta_i(n)\}_n)$ for $i = 1, \ldots, k$ together with a τ-distribution τ. These data can generate various Markov chains such as the augmented chain on $E^k \times \mathcal{I}(k)$ (**7**–3.1, **7**–3.2) or the standard chain on E^k (**7**–4.3). However these chains contain less information than the collection of participators together with τ, for many distinct sets of k participators and choices of τ might give rise to the same chains. Moreover these chains omit the interpretation kernels $\eta_i(n)$ of the participators. For these reasons we equate the level Θ expression of **S** with the (participator, τ)-data even though, by an abuse of language, we sometimes speak of the "dynamical system which expresses **S**" in Θ.

We now consider (ii) of 2.1. Suppose that two complex systems $\mathbf{S}_1$ and $\mathbf{S}_2$ interact and that A_1 and A_2, respectively, are their level Θ expressions as (participator, τ)–ensembles.

$$\mathbf{S}_1\colon\ A_1 = \{(\xi_1,\ \{Q_1(n)\}_n, \{\eta_1(n)\}_n), \ldots, (\xi_k,\ \{Q_k(n)\}_n, \{\eta_k(n)\}_n);\ \tau_k\}$$

$$\mathbf{S}_2\text{:}\quad A_2 = \{(\lambda_1,\, \{R_1(n)\}_n, \{\theta_1(n)\}_n), \ldots, (\lambda_j,\, \{R_j(n)\}_n, \{\theta_j(n)\}_n);\, \tau_j\}. \tag{2.2}$$

We stipulate that a precondition for the interaction is that τ_k and τ_j are compatible (i.e., are part of the same τ-distribution family $\{\tau_i\}_i$). We further stipulate that the interaction itself is expressed in the augmented dynamics of the *joint* participator ensemble. This is the Markov chain on $E^{k+j} \times \mathcal{I}(k+j)$ whose transition probability is

$$\widehat{\langle Q_1(n), \ldots, Q_k(n), R_1(n), \ldots, R_j(n) \rangle}_\tau, \tag{2.3}$$

and whose initial distribution is

$$(\xi_1 \otimes \ldots \xi_k \otimes \lambda_1 \otimes \ldots \otimes \lambda_j)_\tau, \tag{2.4}$$

with the notation of **7**–3.4. Thus the interaction of two systems at level Θ is expressed by the "running" of the participator dynamical chain generated by joining the ensembles representing the two systems separately at that level. Note that such an interaction is meaningful only when both systems employ compatible τ-distributions.

This description of the interaction at a level Θ is natural; it is consistent with the "interactive" character of participators: *any* ensemble of participators subject to the same τ-distribution generates markovian dynamics. It remains to give the specifications (iii) and (iv). For (iii) we must define what it means for two reflexive frameworks Θ and Θ' to be "hierarchically related." The definition must be given in terms of the way in which information flows between the level Θ and level Θ' expressions of a given system. This definition determines the hierarchy, i.e., the ordering of the analytical levels. For (iv) we must specify how information about a level Θ interaction among several systems as stipulated in (ii) is extracted for propagation through the levels of each system. And this specification must comport with the hierarchical relation between levels set forth in (iii).

We may view (iii) and (iv) as imposing constraints on the single-level interaction of (ii). In fact, the information that propagates according to (iii) will be encoded in a form which enables it to pass through the hierarchical connection. (iv) requires that the interaction itself, as specified in (ii), must permit the extraction of this kind of information. This restricts the participator ensembles which may be parties to the interaction. These restrictions constitute the "permissibility condition" on participator ensembles mentioned above, the fulfillment of which is the "form" referred to in (i).

To understand this more concretely, consider systems $\mathbf{S}_1$ and $\mathbf{S}_2$ whose level Θ expressions are as in 2.2, and whose interaction at that level is via the markovian dynamics described in 2.3 and 2.4. In this joint participator dynamics there is no reason why the identity of the original participator ensemble should be retained. By this is meant the following. Suppose that, in isolation, each ensemble has a stable dynamics. When the two ensembles are coupled, their individual stabilities will be disturbed by "cross channelings," i.e., channelings between participators not in the same ensemble. With no constraints on the original systems, we would expect this disturbance to be so great as to eliminate not only the original stabilities but also any possibility of a new pair of stabilities for the individual ensembles. But the interaction data propagated as in (iv) must be meaningful in terms of the individual asymptotics (c.f. Idea 4 of section 1). There must, of course, be a disturbance of these asymptotics in order that an interaction at level Θ of the complex systems $\mathbf{S}_1$ and $\mathbf{S}_2$ can be said to have taken place. But this disturbance should only ***perturb*** the stabilities, not annihilate them. For the individual stabilities are the very grounding of the propagated information.

Thus we suggest that the participator ensembles each need some cohesive stability so that, in this sense, each ensemble maintains its individuality in interaction and so that the resulting perturbations of the dynamics of each ensemble have sufficient regularity to be classified. Assuming this regularity, the perturbation of each system is the interaction data which propagates internally in that system in the sense of (iv). A cohesive stability property which is sufficiently strong in this sense can serve as a "permissibility condition" in 2.1 (i). The enunciation of such cohesive stability properties, and their matching to compatible notions of perturbation regularities, is a central problem in devising models of perceptual hierarchy.

We summarize these ideas in

Terminology 2.5. Let a reflexive framework Θ and a channeling distribution $\tau = \{\tau_k\}_{k=1}^{\infty}$ be fixed. Let $\mathcal{P}$ be a collection whose elements are finite ensembles of participators on Θ.

(i) A ***stability type*** for $\mathcal{P}$ is a class of asymptotic characteristics[1] of the dynamics satisfying the following conditions. The participator dynamics of each

[1] We will not give a precise, general definition of the notion of "asymptotic characteristic." The terminology is intended to include properties of dynamics which can be stated in terms of stationary measures of the dynamics, and, more generally, in terms of "asymptotic" or periodic measures. See, e.g., Revuz

ensemble in $\mathcal{P}$ has asymptotic characteristics in the given class. Moreover, if the dynamics is perturbed by the presence of another ensemble in $\mathcal{P}$ (i.e., when we consider the new dynamics induced on the original ensemble by running the joint participator dynamics generated by it together with another ensemble) then asymptotic characteristics remain in the class (although they may change within it).

(ii) The stability type for $\mathcal{P}$ is said to have a ***perturbation regularity*** if the following condition holds: The variation of the asymptotic characteristics within the class of the stability type, resulting from the perturbations as in (i), has sufficient regularity to be represented in a way which encodes dependency of the variation on the two interacting ensembles in $\mathcal{P}$.

(iii) In the presence of (i) and (ii), we say that $\mathcal{P}$ possesses a ***strong stability type***.

(iv) A ***permissibility condition*** for a stability type is a condition on the ensembles in $\mathcal{P}$, expressible in terms of the action kernels and initial distributions of the constituent participators, which guarantees that the given stability type, with perturbation regularity, will hold for $\mathcal{P}$ (as in (i) and (ii) above). (In other words, it guarantees that $\mathcal{P}$ will have the given strong stability type.

We reiterate that the central idea for propagation between levels is that the information propagated consists in the regular perturbations of an ensemble's asymptotics. Given a framework Θ, the permissibility conditions on ensembles are conditions on the τ-distribution as well as on the data for the constituent participators. (It is expected that the interpretation kernels will play a role in the actual extraction of the data to be propagated—c.f. Idea 2 of section 1.) It seems likely that, even on a given framework, these considerations allow a wide variety of permissibility conditions.

3. Framework Specialization

In this section we discuss more formally how the specialization ideas of the section 1 give rise to canonical schemes for the representation of hierarchical relationships. In the subsequent sections we present two examples of such schemes, the first from computational vision and the second more abstract.

---

chapters 4 and 6.

Terminology 3.1. A *specialization scheme* for a set $\mathcal{P}$ of participator ensembles together with a τ-distribution on a framework Θ, consists of a strong stability type and a corresponding permissibility condition (with the terminology of 2.5).

Intuitively the permissibility condition has two notable consequences. First, the dynamics generated by any ensemble in $\mathcal{P}$ has a stationary measure. Second, the dynamics induced on any such ensemble by running the joint chain generated by it and another ensemble in the set has asymptotic stability which is representable by a measure. The perturbation regularity expresses the relationship that holds, in general, between these latter measures and the original stationary measures. We need not make these intuitions more precise at this point; we only wish to here emphasize that they illustrate an assumption that the specialization scheme is, in some such manner, based on properties of asymptotics which can be expressed in terms of measures.

Any choice of specialization schemes leads to an explicit realization of a hierarchical analytic strategy for participator dynamics which models the stipulations of 2.1. This strategy includes a notion of an information connection between levels of the hierarchy in the sense of 2.1 (iii) and (iv). This connection does not exist between any pair of levels, but only between those which are "hierarchically related": information about perturbation regularities of systems at one level propagates canonically to the next. In this way, we think of the class of levels of the hierarchy, i.e., the class of reflexive frameworks, as having a *relation* defined on it: two frameworks are related if they are connected in this sense. We call the relation *specialization*; each specialization scheme gives rise to a specialization relation.

We now give a formal definition of specialization.

Definition 3.2. Let $\Theta' = (X', Y', E', S', \pi'_\bullet)$ and $\Theta = (X, Y, E, S, \pi_\bullet)$ be reflexive observer frameworks. Let τ be a fixed channeling distribution on Θ, and let a specialization scheme (as in 3.1) be given. Then Θ' is a *specialization* of Θ for τ and for the given specialization scheme if, for some environment $(\mathcal{B}, \Phi)$ supported by Θ' (**5**–2.6), the following hold:

(i)

(a) Let

$$\mathcal{Z} = \{\,(\text{participator}, \tau)\text{-ensembles on } \Theta\}$$

and

$$\bar{\mathcal{Z}} = \{\,(\text{preparticipator}, \tau)\text{-ensembles on } \Theta\}.$$

Let $p : \mathcal{Z} \to \bar{\mathcal{Z}}$ be induced by $(\xi, \{Q(n)\}_n, \{\eta(n)\}_n) \mapsto (\xi, \{Q(n)\}_n)$. Then there are maps $I{:}\, \mathcal{B} \to \mathcal{Z}$ and $\bar{I}{:}\, X' \to \bar{\mathcal{Z}}$ such that $p \circ I = \bar{I} \circ \Phi$. In other words we have a commutative diagram:

$$\begin{array}{ccc} \mathcal{B} & \xrightarrow{I} & \mathcal{Z} \\ \Big\downarrow{\scriptstyle \Phi} & & \Big\downarrow{\scriptstyle p} \\ X' & \xrightarrow{\bar{I}} & \bar{\mathcal{Z}} \end{array}$$

(b) Let D denote the set of those preparticipator (**6**–2.6) ensembles on Θ having the strong stability property of the given specialization scheme. Then $\bar{I}^{-1}(D) = E'$.

(ii)

Let O_1 and O_2 be observers in Θ' (i.e., in $\mathcal{B}$). A channeling between O_1 and O_2 corresponds to the markovian dynamics in Θ resulting from the join of the two participator ensembles $I(O_1)$ and $I(O_2)$ on Θ.

(iii)

(a) The points of Y' parametrize variations of asymptotic characteristics that are meaningful for the preparticipator systems represented (via $\bar{I}$) by the points of X'.

(b) The distinguished premises S' of Θ' correspond to asymptotic variations which express the perturbation regularity provided by the specialization scheme (c.f. 2.5).

(iv)

(a) Given $e' \in E'$ and $x' \in X'$, then $\pi'_{e'}(x') \in Y'$ represents the perturbation of the preparticipator system $\bar{I}(e')$ on Θ which results from its interaction with the system $\bar{I}(x')$.

(b) If $x' \in E'$ then $\pi'_{e'}(x') \in S'$. (This just summarizes the effect of (i)(b) above, i.e., that points of E' correspond to preparticipator ensembles that have the given strong stability type.)

(i)–(iv) of this definition correspond (in toto) to (i)–(iv) of 2.1. The concept of specialization captures the notion of a hierarchical analytic strategy in the form of a relation on the class of reflexive observer frameworks. The environment $(\mathcal{B}, \Phi)$ supported by the framework Θ' (whose existence is required by the definition in order for Θ' to be a specialization) plays only a syntactical role in the definition: the issues which are most central to the question of the specialization of the frameworks themselves are issues of preparticipator dynamics.

Terminology 3.3. *Specialization and Instantiation.* Let Θ' be a specialization of Θ, and let O be an observer in Θ'. If O is a distinguished observer with $\Phi(O) = e' \in E'$ we say that O is the ***specialization*** of the participator system $I(O)$, and that e' is a ***specialization*** of the preparticipator system $\bar{I}(e')$. Similarly, if A is an arbitrary participator (or preparticipator) system on Θ, we say that A ***specializes*** if it has the strong stability property of some specialization scheme. We do not say that $x' \in X'$ is a specialization of $\bar{I}(x')$, or that O is a specialization of $I(O)$ unless x' or O are ***distinguished***, i.e., unless $x' \in E'$ or $\Phi(O) \in E'$.

The term ***instantiation*** denotes the opposite of specialization. For example, with notation as above we say that Θ is an ***instantiation*** of Θ'. However, we use the term "instantiation" to apply to arbitrary (including nondistinguished) configurations and observers, whereas we use the term "specialization" in the distinguished case alone. Thus, for $x' \in X'$, we say that the preparticipator system $\bar{I}(x')$ ***instantiates*** x'; for the observer O in Θ' we say that the participator system $I(O)$ ***instantiates*** O. The maps I and $\bar{I}$ are called ***instantiation maps***. We also say that interactions of participator or preparticipator systems on Θ ***instantiate*** channelings on Θ'. Thus, for observers O_1 and O_2 with $\Phi(O_1) = x'_1$ and $\Phi(O_2) = x'_2$, we say that the markovian dynamics generated by joining the preparticipator ensembles $\bar{I}(x'_1)$ and $\bar{I}(x'_2)$ (or participator ensembles $I(O_1)$ and $I(O_2)$) ***instantiates*** the channeling between O_1 and O_2.

To fix these ideas, let us review how specialized observers make inferences. Let Θ' be a specialization of Θ. Let $e'_1, e'_2 \in E'$, and let $O_{e'_1}$ and $O_{e'_2}$ be observers whose perspectives are $\pi'_{e'_1}$ and $\pi'_{e'_2}$ respectively. Then $O_{e'_1}$ and $O_{e'_2}$ are associated respectively to participator ensembles A_1 and A_2 on Θ. A channeling between $O_{e'_1}$ and $O_{e'_2}$ is instantiated by the participator dynamics generated by joining A_1 and A_2. In the joint dynamics certain properties of the dynamics of the original, separate participator systems are modified, but the systems have sufficient cohesive stability so that these perturbations are not excessively chaotic; the perturbations possess a certain regularity. The distinguished premises S' of the observers in Θ' parametrize structure perturbations with this type of regularity. In particular, the perturbation of the participator dynamics generated in A_1 alone, as a result of A_1 being joined with A_2, corresponds to a point $s' \in S'$. In fact $s' = \pi'_{e'_1}(e'_2)$; it is $O_{e'_1}$'s premise from the channeling between $O_{e'_1}$ and $O_{e'_2}$. Now $O_{e'_1}$ makes an inference from this premise expressed as a conclusion measure, which is a probability measure on ${\pi'_{e'_1}}^{-1}(s') \cap E'$; if η' is $O_{e'_1}$'s interpretation kernel, the measure in question is $\eta'(s', \cdot)$. In terms of the specialization, for each subset C' of E', $\eta'(s', C')$ is the probability that the perturbation represented by s' resulted from joining

A_1 with another participator system which instantiates an element of C'.

So far we have not taken notice of the role of the interpretation kernels in an instantiation. In fact, (i)(a) of 3.2 asserts that the role played by the interpretation kernels of the participators in the ensemble $I(O)$ is relevant to O itself only insofar as O's own interpretation kernel is concerned. Indeed, the environment $(\mathcal{B}, \Phi)$ is not uniquely determined by the definition; essentially distinct choices for $(\mathcal{B}, \Phi)$ correspond to essentially different ways for the interpretation kernel of the observers O in Θ' to relate to the interpretation kernels of the participators in $I(O)$.

The fibre ${\pi'_{e'_1}}^{-1}(s') \cap E'$ contains e'_2, the perspective of the observer which actually channeled with $O_{e'_1}$. But in general there will be many other points in the fibre, and the probability measure will not be concentrated at e'_2; a priori we can say only that the interpretation kernel η' of the specialized observer $O_{e'_1}$ is supported on E'. In fact E' expresses the bias of the specialized observer toward systems with the particular strong stability property specified in the specialization scheme. This means the following. Suppose the instantiation A_1 of $O_{e'_1}$ interacts with *any* participator system, say B, on Θ, in the sense of running the markovian participator dynamics generated by the join of the participator ensembles underlying A_1 and B. Suppose that the resulting perturbation of A_1 exhibits the regularity characteristic of the given specialization scheme, corresponding to a point $s' \in S'$. Then $O_{e'_1}$ will interpret the perturbation as having arisen due to an interaction of A_1 with a preparticipator system on Θ which is the instantiation of some point of E', i.e., a system which has the strong stability property. Thus, in order for $O_{e'_1}$'s inferences to be inductively strong the notion of perturbation regularity which distinguishes the premises S' must be substantially specific to the notion of strong stability which distinguishes the configurations E'. In other words, when a participator system satisfying the permissibility condition undergoes a perturbation with the given regularity, then the chances must be very good that this perturbation was caused by interaction with another permissible participator system.

In the same way we can discuss the instantiation of false objects. As usual, for $e' \in E'$ let $O_{e'}$ denote a distinguished observer in Θ', and let A be the stable participator system in Θ which instantiates $O_{e'}$. Suppose A interacts with an unstable participator system C for which $\bar{I}^{-1}(p(C))$ is in $X' - E'$. Suppose that the resulting perturbation of A exhibits the same regularity property as do perturbations of A resulting from its interaction with stable systems. Then C is an instantiation of a false object for $O_{e'}$. Note that in (iv) of (3.2) no stipulation is made about the premises of *nondistinguished* observers in Θ' which result from channelings with any other observer, distinguished or nondistinguished. This is so even though for nondistinguished as well as distinguished observers channelings are instantiated in the same way, namely by the interaction of the participator systems in Θ which correspond to the observers' configurations in X'. The difference is that for a nondistinguished observer in a reflexive framework there is no a priori relation between its configuration and its perspective map; thus, even though a channeling for a nondistinguished observer arises from an interaction of the instantiating participator system associated to the observer's configuration, and even though the premise resulting from the channeling is a point of Y' corresponding to a structure perturbation which is in principle meaningful for the participator system, yet in the absence of any information about the perspective map there is no basis from which to impute meaning to the premise of the nondistinguished observer ***in terms of the interaction***.

Given that the permissibility condition and the perturbation regularity of a specialization scheme depend on stationary or asymptotic measures, it follows that the role of true perception in specialization is twofold. First, the existence of true perception is a step in the direction of stability in the sense that true perception requires stationary measures. Of course, the strong stability needed for specialization requires more than the simple existence of stationary or asymptotic measures for each instantiating system. For instance, these systems need to stabilize in the presence of other such systems in some way yet to be defined. Second, true perception—be it on the part of all or merely some of the participators in the instantiating system—is necessary for the conclusions of the specialized observer to be inductively strong. In fact, the distinguished specialized observer O infers the identity of the system which interacts with its instantiation $I(O)$; the premise for this inference is the perturbation of $I(O)$'s structure which results from the interaction.

Recall that 3.2 (i) states that in a given environment $(\mathcal{B}, \Phi)$ the interpretation kernels of the participators in the ensemble $I(O)$ functionally constrain the interpretation kernel of O itself. However, the definition does not stipulate any details about this constraint: the manner in which the specialized observers' interpretation kernels are related to those of the participators in the instantiations is a "free variable" in the specialization relation between frameworks. The various choices correspond to the various environments $(\mathcal{B}, \Phi)$ which fit in the definition 3.2. In particular there are many possibilities for formulating interpretation strategies for the specialized observers, whose principle is to exploit in some manner true perception down at the level of the instantiation. And it is such strategies which intuitively lie at the heart of the specialization idea.

There is not a unique way to specialize, nor to instantiate, a given framework. Beginning with the framework Θ we can consider various specialization schemes which make sense for Θ. But even if we fix the specialization scheme there is not a unique framework which is a specialization of Θ. For example one can restrict attention to various subclasses of all those participator systems which specialize in the sense of the given scheme, and then consider a framework Θ' whose distinguished configuration set E' parametrizes the participator ensembles in one such subclass. The parametrization itself can be made in various ways. Once this is done, however, the perspective maps $\pi'_{e'}$ in Θ' are essentially determined by the specialization scheme. In fact, $\pi'_{e'_1}(e')$ is the point of S' which represents the dynamical perturbation of the participator system on Θ which instantiates e'_1, resulting from the join of that system with the system which instantiates e'.

As we have remarked above, the concept of specialization of frameworks defines a relation on the set of all reflexive frameworks which we think of as a hierarchy relation $\dashv$. We do not prove here that specialization is transitive, but a few considerations make this plausible. Denoting specialization by $\dashv$, if $A \dashv B$ then the premises of A are perturbations of stationary measures for the dynamics of ensembles on B. If $B \dashv C$ then premises of B are perturbations of stationary measures for ensembles on C. For transitivity $A \dashv C$ must also be true; the premises of A must be perturbations of stationary measures for ensembles on C as well as on B. But, since $B \dashv C$, configurations (elements of X) of B correspond to ensembles on C. And at each instant each participator in an ensemble on B must manifest as a configuration of B, i.e., (via the map $\bar{I}$) as an ensemble on C. Thus, ensembles on B can be thought of as ensembles of ensembles on C. It is then at least plausible that the premises of A could correspond to perturbations of the stationary measures of ensembles on C.

We notice that for one framework to be a specialization of another does not imply that the intrinsic mathematical properties of the frameworks are different. For example, it is possible that two frameworks are abstractly isomorphic, yet one is a specialization of the other. Thus specialization provides a universal way to interpret frameworks in terms of others via the relation in the lattice, but does not constrain the intrinsic, purely mathematical, structure of the individual frameworks.

We have not discussed the way in which information propagates downward in the lattice, only upward. The downward propagation has to do with the effect that the presence of specialized systems have at the lower level. Intuitively, they propagate coherence. However, their effect (if one looks at dynamics down in Θ which are really joint dynamics with the specialized system, but are represented as though the specialized system is not there) may be described as a modification of the τ-distribution or of the action kernels in Θ. These two formulations of their effect may be equivalent, and the expression of that equivalence may be a “natural law,” like Newton’s law relating force and acceleration or more probably like the Einsteinian version relating metric geometry and force-acceleration. For remember that the τ-distribution is somehow intimately related to metric-like notions on E, while changes in the action kernels of participators is intimately related to "acceleration" in the same spirit in which the action kernels themselves correspond to "velocity."

4. On Ullman's incremental rigidity procedure

4.1. Preliminary remarks and overview. We present an example of specialization inspired by Ullman's "incremental rigidity scheme," a procedure whereby a viewer can generate and update an internal three-dimensional model of an external object as the object moves in space relative to the viewer. One assumes that the object consists of, say, $n + 1$ feature points and that the "correspondence problem" has been solved, i.e., that the viewer can track each point over time. We further assume that the viewer deploys a moving coordinate system in which the same one of these points is always at the origin. Then the vectors from this origin to the other n points describe at each instant of time. Finally we assume that the viewer has access only to two-dimensional orthographic projections (onto some fixed image plane) of these n vectors. The viewer updates its internal three-dimensional model based on

(i) its current model,
(ii) the latest two-dimensional projection of the object.

The viewer chooses that new model, from among all those compatible with the new information (ii), whose three-dimensional structure differs minimally from that of the current model. If the resulting sequence of models converges to a stable rigid structure then the viewer infers that the object has that same rigid structure. If, in the limit, the sequence of models exhibits some periodicity, then the viewer infers that the object has the type of quasi-rigidity expressed by the periodicity.

Ullman called this the "incremental recovery of 3-D structure from rigid and rubbery motion." The phrase "recovery of 3-D structure" here refers to the conclusion of an inference about the stable three-dimensional structure of the object, not about its instantaneous three-dimensional structure. One way an object might exhibit a stable or long-term 3-D structure is to forever move rigidly. Another way is to expand and contract periodically.

Just as the conclusion of the inference in Ullman's scheme refers to stability of structure, so also the premise of the inference depends upon a form of stability. An essential feature of Ullman's scheme is that the premise of the inference is derived from the long-term, i.e., asymptotic, behavior of a certain dynamical interaction. For Ullman this is an interaction between viewer and physical object. For us all interactions are between observers; physical objects represent the conclusions observers reach in consequence of their interactions.

We now study an observer-theoretic treatment of this inference. We consider a symmetric observer framework Θ in which the observers' inferences regard *instantaneous* 3-D structure. On this framework we have a participator dynamics whose asymptotic stabilities give rise to premises for a "higher level" observer which infers a *long-term* structural regularity. This is a specialized observer, i.e., an observer in a framework Θ' which is a specialization of Θ in the sense of section three. Thus the observers in Θ' infer long-term stabilities; the observers in Θ infer instantaneous rigidity. Now, neglecting translation, instantaneous rigid motion is the same as instantaneous rotation, so we will take Θ to be the symmetric framework of instantaneous rotation observers studied in **5**–6. Recall that a distinguished premise in this framework consists of two frames of n vectors, together with a reference axis, which are compatible with an interpretation that the frames are related by a rotation of $\mathbf{R}^3$ about that axis. In practice this means that, in our incremental rigidity procedure, two such consecutive frames of n vectors are required to trigger a step. This is in contradistinction to Ullman's original procedure, where any single frame of n vectors triggers a step.

We begin with the symmetric framework $\Theta = (X, Y, E, S, G, J, \pi)$ of instantaneous rotation observers. We will describe a specialization scheme and a framework $\Theta' = (X', Y', E', S', \pi'_\bullet)$ which is a specialization of Θ for this scheme. This specialization is simple. Points of E' correspond (via $\bar{I}$ of 3.2, (i)) to preparticipator ensembles on Θ consisting of *one* preparticipator. Similarly, the distinguished observers in Θ' correspond (via I of 3.2, (i)) to participator ensembles consisting of one participator. For example, let O' be a distinguished observer in Θ' whose configuration $\Phi(O') \in E'$ corresponds to the ensemble consisting of the sole preparticipator A in Θ; we say "A instantiates O'." O' uses the incremental procedure to make inferences as follows. Suppose that A is involved in a participator dynamics on Θ with another participator B (or more generally some set of participators). The asymptotic behavior of this dynamical interaction instantiates a single channeling at the level of Θ' for O'. From this channeling O' infers, if possible, B's rigid or quasi-rigid structure. Here is how we think of this as an incremental rigidity scheme:

(i) At any time t (in the reference time for the dynamics in Θ) the state $e(t) \in E$ of A is the "current model" of the instantaneous structure of B.

(ii) A's action kernel is defined such that A executes the updating procedure associated with the scheme.

If this dynamics induces the right kind of asymptotic regularity on the trajectories of A, then O' infers that B has the appropriate stability. The existence of such an asymptotic regularity on A's trajectories corresponds to what Ullman calls the "convergence" of the incremental procedure. In our terminology it means that O''s premise resulting from the channeling is distinguished. In (i) above we used the quotes on "current model" to stress that it has no a priori perceptual status, even instantaneously, at the level of the specialized observer O'. Indeed, an instant of time for O' is that time in which a channeling occurs for O'; and this must correspond to sufficient time at the level of Θ for the entire participator dynamics involving A to reveal asymptotic stability. Thus an instant of reference time on Θ is not meaningful for O'. With this in mind we can present the situation in more detail.

4.2. We use the notation and terminology of **5**–6 for the framework Θ of instantaneous rotation observers. Let us fix the point $c_0 \in E$, and henceforth denote the fundamental map π^{c_0} (**5**-6.20) simply by π. For example, for convenience of visualization we can take c_0 to be a configuration whose reference axis $\mathcal{A}$ is the positive z-axis, and whose **v** is the unit vector in the positive x direction.

We now discuss A's action kernel $\{Q_e\}_{e \in E}$. Recall that this is a family of markovian kernels on E, one for each $e \in E$ (**7**–1.1). The kernel Q_e describes how A moves in response to a channeling when A is at e. In our case the action kernel will be symmetric, i.e., $\{Q_e\}_{e \in E}$ is generated by a single markovian kernel $Q\colon J \times \mathcal{J} \to [0, 1]$. Given Q, we define Q_e by $Q_e(e_1, \Delta) = Q(e_1 e^{-1}, \Delta e^{-1})$. This is the probability that A will move from e into the set $\Delta \subset E$ given that it received a channeling from e_1. The fact that the action kernel is symmetric means that this probability depends only on the position of e_1 relative to e, and of Δ relative to e (in the sense of the group action of J on E). Finally, we recall that $Q(\jmath, \cdot) = Q(\jmath', \cdot)$ if $\pi(\jmath) = \pi(\jmath')$.

Suppose that, at a particular time t, A is at e and A channels with another participator at e_1. This channeling results for A in the observation event $s = \pi(e_1 e^{-1}) \in S$. The updating procedure of (ii) above means, firstly, that A then moves so that its new state is a possible state of the participator which just channeled to him, i.e., A's new state will lie in $\pi^{-1}(s)$. Secondly, it means that the new state selected in $\pi^{-1}(s)$ will minimize the distortion of the underlying rigid structure entailed in the state change.

A's motion, then, is based on minimizing a certain nonnegative function ϕ on $\pi^{-1}(s)$, a function which measures the structural modification associated to the move. Now everything is already relativized with respect to A's perspective e; if $\jmath \in \pi^{-1}(s)$, the selection of $\jmath$ means that A will move from e to $\jmath e$. Thus the function in question is naturally a function on J, because the elements of J

are intrinsically the "moves" whose structural effect we wish to measure.[2] We would expect, then, that the definition of the function ϕ itself is independent of both e and s; then, whenever a premise s is presented, the procedure is to minimize ϕ on the subset $\pi^{-1}(s) \subset J$. $\pi^{-1}(s)$ is a one-dimensional manifold with four connected components; this follows from the fact that the same is true for $p^{-1}(s)$ and that $\pi = f_{c_0} \circ p$ (**5**–6.20) where f_{c_0} is an isomorphism.

The function ϕ is by no means uniquely specified, for one may conceive of many different ways of testing "rigidity." However, the representation **5**–6.17 of J leads naturally to a description of a class of reasonable ϕ's. In fact, in terms of this representation, and of the expression for $\jmath e$ in **5**–6.19, it is easy to see that it is precisely the nonzero γ_i's, ζ_i's and λ_i's which contribute nonrigidity to the transformation $e \mapsto \jmath e$. δ simply augments the magnitude of the angular velocity of the instantaneous rotation embodied in e; β rotates the entire structure e. More specifically, the α's and ζ's perturb the structure additively while the λ's perturb it multiplicatively. Hence, we should require that

4.3. ϕ is a monotone function of $|\gamma_i|$, $|\zeta_i|$, and $|\lambda_{i-1}|$, where $|\gamma_i|$ denotes the distance (along the circumference) from γ_i to the identity element of the circle group $\mathbf{S}^1$.

Now given ϕ we can use it to define the action kernel Q of A. Intuitively, we want to minimize ϕ on $\pi^{-1}(s)$, and then let $Q(\jmath, \cdot)$ be Dirac measure concentrated at the minimum (when $\pi(\jmath) = s$). This is a deterministic action kernel; A's next state is uniquely determined by its current state and the observation event s which results from the channeling. But in general ϕ has a no unique minimum on each $\pi^{-1}(s)$. Therefore we consider nondeterministic action kernels for A. And we need not minimize ϕ. Instead we proceed as follows: Let μ denote some natural "unbiased" measure (such as Haar measure) on J and let

$$\mathcal{F} = \{\phi \text{ satisfying } (4.3) | \int_E \frac{1}{\phi} d\mu = 1\}. \tag{4.4}$$

[2] The identification of J with E simply gives a way to "visualize" the elements of J. In this sense the choice of c_0 in the definition of π means that A "thinks of itself" as c_0, and refers to an element of $j \in J$ in terms of what A would then become if it were modified by j (see **5**–6.22 ff).

We identify functions in $\mathcal{F}$ that differ on a set of μ measure zero, i.e., we think of $1/\phi$ as an element of $L_1(J,\mu)$. We can then identify $\phi \in \mathcal{F}$ with the probability measure $\mu_\phi = (1/\phi)d\mu$ on J. Now there is a canonical way to generate the kernels Q given ϕ's:

4.5. To $\phi \in \mathcal{F}$ we associate the kernel defined by $Q(\jmath,\cdot) = m_\pi^{\mu_\phi}(\pi(\jmath),\cdot)$.

That is, Q is the rcpd of μ_ϕ with respect to π. (If we wish, we can replace $\mathcal{F}$ by a suitable completion. Then among the new limit measures we recover the Dirac measures of the deterministic case mentioned above.)

In this way any $\phi \in \mathcal{F}$ is associated with *an* incremental rigidity procedure, the one executed by the participator A whose action kernel is defined as in 4.5. Intuitively, if A interacts with a participator B then A converges asymptotically to the trajectory of B. The question of whether or not there is convergence in any particular case depends a priori on the choice of ϕ, on the initial distribution of A, and on the motion and shape of B. We do not consider this question in detail. The point of view we want to emphasize here is that of the "rigid object" as a *conclusion* of a specialized observer, not as an *object of perception* for that observer.

4.6. If $\{T_e\}_{e\in E}$ is B's action kernel then, for any e and e_1, the measures $T_e(e_1,\cdot)$ are supported on the orbit through the point e of the subgroup R of J given by $\alpha_j = 0$, $\zeta_j = 0$, $\lambda_j = 1$ (for all j). This subgroup of J, parametrized by β and δ, is isomorphic to $\mathrm{SO}(3,\mathbf{R}) \times S^1$. Thus B will stay in a fixed R-orbit in any interaction.

There is another natural way to think about the ϕ's in terms of this subgroup R: each choice of ϕ as in 4.3 gives a "distance function" to R on J. To see what this means in terms of participator dynamics on E, consider a participator A on Θ whose action kernel Q is of the form of 4.5 for such a distance function ϕ. Suppose that at time t (reference time on Θ) A is at $e \in E$ and channels with an observer at $e_1 \in E$. Suppose $e_1 = j_1 e$, $j_1 \in J$. A then moves to je, where j is in the fibre $\pi^{-1}(\pi(j_1))$, with a probability that depends on the distance of j to R; the smaller the distance, the greater the probability. Thus, the effect of the action kernel Q is to make A tend to move on R-orbits in E.

We now give a sample definition for the specialized framework Θ', and for the specialization scheme that leads to it. We start with the specialization scheme; we use the terminology of 3.1. Let R be the subgroup of J defined in 4.6. The strong stability condition on participator ensembles is the following *asymptotic R-orbit property*: the dynamics admits a stationary measure which is supported, say, on a finite union of R orbits in E^k, where k is the number of participators in the ensemble. Recall that the strong stability condition must hold not only for the dynamics of each permissible participator ensemble individually, but must also hold for the dynamics induced on it by the joint system it generates with any other permissible ensemble. In our case this is part of the definition of the condition. Thus the strong stability condition is really a condition on *sets* of ensembles, not just on individual ensembles: any condition which defines a set of participator ensembles with these properties can serve as a "permissibility condition." The perturbation regularity is that the R orbit property of the asymptotics is preserved under perturbation, i.e., under interaction with another permissible ensemble.

We now describe one possible specialized framework Θ' for this specialization scheme, one which is especially (and artificially) simple. We assume that we have a fixed τ-distribution on Θ. We can let X' be a set of (preparticipator, τ)-ensembles each consisting of only one preparticipator (and the τ is the fixed one); the map $\bar{I}$ of 3.2 is then just the inclusion map. The elements of E' involve preparticipators whose action kernel is like the one given in 4.5 for a particular choice of ϕ. We assume that the functions ϕ and the initial measures of these preparticipators have been chosen so that the set E' has the following property:

4.7. The dynamics generated by a preparticipator in E' with any other preparticipator in X' has a stationary measure in E^2; the dynamics generated by two preparticipators in E' has a stationary measure supported on a finite union of R-orbits in E^2.

By saying that the dynamics "has" a stationary measure we mean that the initial measure converges to the stationary measure under the action of the dynamics. Also, when we say a "preparticipator in E'" we mean the ensemble in E' consisting of that one preparticipator. It may require work to show that there exist ϕ's and initial measures such that the resulting E' has this property. However, since our objective here is just to illustrate the basic ideas of specialization, we simply assume they exist.

Let $pr_1 \colon E^2 \to E$ be projection on the first factor. Let Y'' denote the set of measures on E^2 which are stationary measures of the joint dynamics generated by some $e' \in E'$ and some $x' \in X'$; we here use the property of E' in italics in the preceding paragraph. Now let Y' denote the set of all measures on E which are of the form $pr_{1*}(\rho)$ for some $\rho \in Y''$. Let S' denote those measures in Y' which arise as above in the case where both e' and x' are in E'. Note that if a measure ρ on E^2 is supported on a finite union of R-orbits in E^2, then $pr_{1*}(\rho)$ is supported on a finite union of R-orbits in E. It follows that each measure in S' is supported on a finite union of R-orbits in E. We can now define $\pi'_{e'}(x')$: it is the element of Y' which represents $pr_{1*}(\rho)$, where ρ is the stationary measure on E^2 of the joint dynamics generated by e' and x'. We have thus defined the reflexive framework $\Theta' = (X', Y', E', S', \pi'_\bullet)$; note that we have not shown this framework to be ***symmetric***. Conditions (ii), (iii), (iv) of Definition 3.2 are satisfied for Θ' with respect to our given specialization scheme. And the map $\bar{I}$ of (i) of 3.2 is defined as the inclusion. We have not yet discussed the map I of 3.2(i) (and the significance of the commutative diagram there) for our present situation; we consider this briefly below.

We discuss the relevance of Θ' to the original problem of rigid object perception. The elements of E' do not represent rigid objects, because the action kernels of the preparticipators of E' are of the type of the Q of 4.5 for some ϕ, and not of the type of the T of 4.6. In other words, unlike a "rigid object," a preparticipator with an action kernel Q does not remain in a fixed R-orbit regardless of its channeling interactions. It is still possible that some elements of X' represent rigid objects since for such elements not in E' we have made no stipulation about the action kernel of the preparticipator. We regard a "rigid object" as being a ***conclusion*** of a specialized observer. In fact, it is the conclusion of a distinguished observer in Θ' resulting from a premise $s' \in S'$ which is a measure (a $pr_{1*}(\rho)$ as above) supported on a ***single*** R-orbit. In general, a point of S' is a measure supported on a finite union of such orbits; the conclusion resulting from such a premise is a "quasi-rigid" object which is a superposition of "rigid conclusions." These latter correspond to the components of the measure on the distinct orbits of the union. If O' is a distinguished observer in Θ' whose configuration is e' then the conclusion of O' in response to the premise s' is a probability measure on ${\pi'_{e'}}^{-1}(s')$; in fact it is the measure $\eta'(s', \cdot)$, where η' is the interpretation kernel of O'. The rigid (or quasi-rigid) object is O''s representation of this measure.

The definition of specialization (3.2) requires that we adduce a particular environment $(\mathcal{B}, \Phi)$ supported by Θ'. Then when we speak of an "observer in Θ'" having a property which shows some aspect of the specialization we mean an observer in this $\mathcal{B}$. To define an environment on Θ', or at least to define the distinguished part of it, we describe the interpretation kernels which are associated with various points of E'. The commutative diagram of (i) of 3.2

means that, for the environment $(\mathcal{B}, \Phi)$, there is some relationship between the interpretation kernels of the observers O' in $\mathcal{B}$ and the interpretation kernels of the participators in the ensemble $I(O')$. For example, consider the observer O' whose configuration $\Phi(O')$ is $e' \in E'$. e' is an ensemble consisting of a preparticipator (ξ, Q) on Θ. The commutative diagram then requires that O' itself is associated (via I) to an ensemble consisting of the one participator $A = (\xi, Q, \eta)$, for some η. Let η' denote the interpretation kernel of O'. A complete demonstration that Θ' is a specialization of Θ requires that we state a relationship between η' and η which holds for all the observers in a set $\mathcal{B}$. We will not analyze this further here; we will only reiterate the basic idea that *true perception* plays a major part in this relationship. Namely, the assumption that η truly reflects the asymptotic behavior of participator A alone is the basis of a strategy expressed by η', a strategy for the specialized perceiver O' to make inferences based on perturbations of those asymptotics.

5. Chain-bundle specialization

We now sketch one approach to specialization, called "chain bundle specialization," which can be applied to symmetric observer frameworks under certain conditions. Starting with a symmetric framework $\Theta = (X, Y, E, S, G, J, \pi)$, we use a specialization scheme (3.1) which exploits the group action of J on E to define the permissibility condition on participator ensembles *and* of the perturbation regularity. The scheme is valid under conditions which we make explicit below. The mathematical content of certain of these conditions (which pertain to the perturbation regularity) is not yet clarified; for this reason the approach is speculative. However, we believe that the scheme is valid for natural and nontrivial classes of examples; we discuss this after presenting more details.

5.1. We introduce notation for certain elementary constructions associated with measurable group actions. Let Γ be a measurable group and Z a measurable space; let a measurable left action $z \to \gamma z$ of Γ on Z be given. Then there is an induced left action of Γ on the set $\mathcal{Z}$ of measurable functions on Z, namely

$$f \to \gamma f, \quad (\gamma f)(z) = f(\gamma^{-1} z).$$

This action is linear, i.e., $\gamma(f_1+f_2)=\gamma f_1+\gamma f_2$. We also use the notation ${}^{\gamma}f$ in place of γf. Thus we will also write the action as

$$f \to {}^{\gamma}f.$$

In this manner we think of $\gamma \in \Gamma$ as a linear operator on $\mathcal{Z}$ or on the space $\mathcal{Z}_b$ of bounded measurable functions. Now let K be a kernel on Z. K may be viewed as a linear operator on $\mathcal{Z}$ via

$$Kf(z) = \int K(z,du)f(u).$$

We can then define a left action of Γ on kernels by

$$K \to {}^{\gamma}K = \gamma K \gamma^{-1};$$

the notation on the right means $\gamma \circ K \circ \gamma^{-1}$ in the sense of composition of linear operators on $\mathcal{Z}$. The notation ${}^{\gamma}K$ thus makes sense for *any* operator on $\mathcal{Z}$ (not only those associated to kernels). If K preserves bounded functions then so does ${}^{\gamma}K$. In terms of arguments, we have explicitly

$${}^{\gamma}K(z,A) = K(\gamma^{-1}z, \gamma^{-1}A).$$

where A is a measurable set in Z; this is easily checked.

Any measure μ on Z can be viewed as a linear functional on $\mathcal{Z}$. In this sense, for $\gamma \in \Gamma$ we can define ${}^{\gamma}\mu$ to be the composition $\mu \circ \gamma^{-1}$. This gives a left action of Γ on the space $\mathcal{M}$ of measures on Z:

$$\mu \to {}^{\gamma}\mu.$$

$${}^{\gamma}\mu(A) = \mu(\gamma^{-1}A).$$

Proposition 5.2. With the notation as above,

1. For any operator K and function f,

$${}^{\gamma}(Kf) = {}^{\gamma}K\, {}^{\gamma}f$$

2. If K is a kernel and μ is a stationary measure for K:

$$\mu K = \mu \Rightarrow {}^{\gamma}\mu\, {}^{\gamma}K = {}^{\gamma}\mu.$$

Proof. Straightforward.

Definition 5.3. Let Γ be a measurable group, and suppose Z and W are spaces on which Γ acts measurably. Let $p\colon Z \to W$ be a measurable map. p is called a Γ-*homomorphism* if $\gamma p(z) = p(\gamma z)$ for all $\gamma \in \Gamma$ and $z \in Z$. In this case the data

$$\begin{array}{c} Z \\ \big\downarrow p \\ W \end{array}$$

is called a Γ-*bundle* if p is surjective. If the action of Γ on Z (and hence on W) is transitive, it is called a ***transitive*** Γ-***bundle.*** Z is called the "total space" of the bundle, p is called the "projection map," and W is called the "base space."

We will build bundles from participator systems on a given symmetric framework Θ. Under certain conditions we will be able to view the total space, base space, and projection map of the bundle as the distinguished configuration space, the distinguished premise space, and the perspective map of a new symmetric framework Θ' which is a specialization of Θ.

5.4. We begin with a symmetric observer framework $\Theta = (X, Y, E, S, G, J, \pi)$ with fixed τ-distribution τ. Let $k > 0$ be an integer, and consider k symmetric action kernels $Q_1, ..., Q_k$ on Θ. We can then construct the markovian kernels $\widehat{P_0} = \widehat{\langle Q_1, \ldots, Q_k \rangle_\tau}$ on $E^k \times \mathcal{I}(k)$, and $P_0 = \langle Q_1, \ldots, Q_k \rangle_\tau$ on E^k. $\widehat{P_0}$ and P_0 are, respectively, the transition probabilities for the augmented and standard dynamical Markov chains respectively, of an ensemble of k kinematical (i.e., time homogeneous) participators whose action kernels are $Q_1, ..., Q_k$. Now the properties of participator ensembles which are relevant to a specialization scheme may be best expressed in terms of the augmented dynamics of the ensemble, rather than the standard dynamics. Nevertheless for simplicity of exposition we restrict our attention to the standard dynamics.

The group J of the framework Θ acts measurably on E^k on the left via its given measurable action on E:

$$j(e_1, ..., e_k) = (je_1, ..., je_k).$$

More generally, let J' be a group which is a ***measurable extension of a subgroup*** of J. This means that we are given a group homomorphism $\alpha\colon J' \to L$ where $L \subset J$ is a subgroup; further, J', L, and α are measurable. In this case the action of J on E^k induces a measurable left action of J' on E^k by letting $\gamma e = \alpha(\gamma)e$ for $\gamma \in J', e \in E^k$.

Assume we are given such a J', which we view as acting on E^k in this manner. Suppose ν_0 is a stationary measure for the kernel P_0 on E^k. Then we can define ${}^{\gamma}P_0$ and ${}^{\gamma}\nu_0$ as in 5.1, and the conclusions of Proposition 5.2 hold, namely

5.5. For all $\gamma \in J'$, and measurable functions f on E^k,

$$ {}^{\gamma}(P_0 f) = {}^{\gamma}P_0 \, {}^{\gamma}f, $$

and ${}^{\gamma}\nu_0$ is a stationary measure for ${}^{\gamma}P_0$:

$$ {}^{\gamma}P_0 \, {}^{\gamma}\nu_0 = {}^{\gamma}\nu_0. $$

Now we can describe our chain bundle. Let

$$ E_1' = \{({}^{\gamma}P_0, \ {}^{\gamma}\nu_0) \mid \ \gamma \in J'\}, $$

$$ S' = \{{}^{\gamma}\nu_0 \mid \ \gamma \in J'\}, $$

$$ \pi_1'\colon E_1' \to S', \quad \pi_1'(P,\nu) = \nu. \tag{5.6} $$

The left action of J' on kernels and on measures gives a left action of J' on E_1', namely

$$ \gamma_1({}^{\gamma}P_0, \ {}^{\gamma}\nu_0) = ({}^{\gamma_1}({}^{\gamma}P_0), \ {}^{\gamma_1}({}^{\gamma}\nu_0)) = ({}^{\gamma_1\gamma}P_0, \ {}^{\gamma_1\gamma}\nu_0). $$

It is then clear that

$$ \begin{array}{c} E_1' \\ \big\downarrow {\scriptstyle \pi_1'} \\ S' \end{array} $$

is a ***transitive*** J'***–bundle.***

The terminology “chain bundle” indicates that points in the total space E_1' are γ-homogeneous Markov chains on E^k. The chain is specified by its transition probability, namely ${}^{\gamma}P_0$ for some $\gamma \in J'$, and its starting measure ${}^{\gamma}\nu_0$. This starting measure is also a stationary measure for the chain by 5.5 since, by hypothesis, ν_0 is stationary for P_0. For $\nu \in S'$, ${\pi_1'}^{-1}(\nu)$ is the subset of E_1' consisting of all those chains whose specified starting (and stationary) measure is ν.

Consider a preobserver $O' = (X', Y', E_1', S', \pi_1')$, where X' is a set of Markov chains on E^k containing E_1', and Y' is a set of measures on E^k containing S'. The inferences of O' are at a “higher level” than the inferences of observers in Θ: each premise of O' represents a possible stability of a whole dynamical system in Θ, and a corresponding conclusion represents a Markov chain in E^k with that stability. This description of the meaning of O''s observations obtains because the group action on E_1' preserves stationarity of the measures, as in 5.5. However, the inferences of O' are not even *ascendants* of those of Θ. The reason for this is that while the initial Markov chain (P_0, ν_0) is a participator chain in the sense that $P_0 = \langle Q_1, ..., Q_k \rangle_\tau$ for some action kernels $Q_1, ..., Q_k$, the same is not necessarily true for $({}^{\gamma}P_0,\ {}^{\gamma}\nu_0)$ for arbitrary $\gamma \in J'$. This means that the premise ${}^{\gamma}\nu_0 \in S'$, while it is a stationary measure for *some* markovian dynamics in E^k, is in no meaningful way derived from conclusions of observers in Θ.

On the other hand, suppose

Assumption 5.7. For every $\gamma \in J'$ we can find action kernels ${}^{(\gamma)}R_1, ...,\ {}^{(\gamma)}R_k$ on Θ such that

$$ {}^{\gamma}P_0 = \langle\ {}^{(\gamma)}R_1, ...,\ {}^{(\gamma)}R_k \rangle_\tau . $$

(We may suppose that when γ is the identity element of J' the ${}^{(\gamma)}R_i$'s are the original Q_i's). Then for each γ we can imagine an ensemble ${}^{(\gamma)}A$ of kinematical participators,

$$ {}^{(\gamma)}A = \{\ {}^{(\gamma)}A_i\}_{i=1}^{k}, \quad {}^{(\gamma)}A_i = (\xi_i,\ {}^{(\gamma)}R_i, \eta_i) \tag{5.8} $$

where the starting measure $\xi_1 \otimes ... \otimes \xi_k$, together with the transition probability ${}^{\gamma}P_0$, give rise to a chain on E^k with stationary measure ${}^{\gamma}\nu_0$. Let us also suppose that these participators have stably true perception (**8**–5.8), so that the interpretation kernels η_i are related to the stationary measure ${}^{\gamma}\nu_0$, via an rcpd construction similar to the one used with the “$\mathcal{D}$ operation” of **8**–5.4 ff. We may even imagine that the situation at hand is sufficiently constrained so that the collection of η_i's is informationally equivalent to the stationary measure ${}^{\gamma}\nu_0$. Under these conditions we can say that the premises in S'—the various measures ${}^{\gamma}\nu_0$—are deduced from the conclusions of observers in Θ, namely the observer manifestations of all the participators ${}^{(\gamma)}A_i$ for $\gamma \in J'$. It follows that, in this context, elements of S' represent premises of inferences which are ***ascendants*** of the inferences in Θ.

Thus, assuming 5.7 and 5.8, let

$$E' = \{\ {}^{(\gamma)}A \mid \quad \gamma \in J'\}.$$

$$\pi_0'\colon E' \to S',\ \pi_0'(\ {}^{(\gamma)}A) =\ {}^{\gamma}\nu_0. \tag{5.9}$$

J' acts on E' simply by acting from the left on the symbol ${}^{(\gamma)}$ in ${}^{(\gamma)}A$. In this way we can consider E' and E_1' to be isomorphic as measurable J'–spaces, and then the J'–bundles $\pi_0'\colon E' \to S'$ and $\pi_1'\colon E_1' \to S'$ are isomorphic. But in contrast to the inferences from S' to E_1', the inferences from S' to E' are now ascendants of inferences in Θ. And what is more, they have a chance to be inferences of observers in a ***specialization*** of Θ; for the configuration space of such a specialization consists by definition of participator ensembles in Θ. In other words, assuming 5.7 holds, we may be able to construct a specialization Θ' of Θ in which E' and S' are the distinguished configuration and premise spaces; we indicate, however, that 5.7 alone is not sufficient for the existence of Θ'. We will discuss this question below, but first we present an important class of examples where 5.7 holds.

The action of J' on E' described above is not intrinsic; it has been transported artificially to E'. We have indicated this by writing the superscript γ in parentheses in ${}^{(\gamma)}A$ and ${}^{(\gamma)}R_i$. The point is that these superscripts do not here refer to any well-defined mathematical operation, as they do in the case of the ${}^{\gamma}P_0$. In effect, in 5.7 we assume only that for each $\gamma \in J'$ an ${}^{(\gamma)}A$ exists; we have not assumed that the ${}^{(\gamma)}A$ are generated by some intrinsically defined group action on participator ensembles, starting from some such ensemble in which the action kernels are the original $Q_1, ..., Q_k$. However in the class of examples we now present there is such an intrinsic action which generates the ${}^{(\gamma)}A$.

Recall (5.4 ff) that we are starting with a group J' which is an extension of a subgroup L of J; J is the distinguished structure group of our original framework $\Theta = (X, Y, E, S, G, J, \pi)$.

Proposition 5.10. Suppose that (1) τ is a translation-invariant τ-distribution on Θ, and that (2)

$$\begin{array}{c} J \\ \Big\downarrow{\scriptstyle \pi|_J} \\ S \end{array}$$

is a bundle for the action of L on J by conjugation. Then there is a left action of J' on the set of symmetric action kernels on Θ; in terms of the generator Q of the action kernel (**7**–1.1) the J'-action is expressed by

$$Q(\cdot, \bullet) \to {}^{(\gamma)}Q(\cdot, \bullet) =_{def} \; Q(\gamma^{-1} \cdot \gamma, \gamma^{-1} \bullet \gamma), \; \gamma \in J'. \tag{5.11}$$

with the property: if $Q_1, ..., Q_k$ are any symmetric action kernel generators on Θ, then

$${}^{\gamma}\langle Q_1, ..., Q_k\rangle_\tau = \langle \, {}^{(\gamma)}Q_1, ..., \; {}^{(\gamma)}Q_k\rangle_\tau. \tag{5.12}$$

Proof. To say that Q is the generator of a symmetric action kernel on Θ means that $Q\colon J \times \mathcal{J} \to [0,1]$ is a kernel with the property that if $\pi(j_1) = \pi(j_2)$ then $Q(j_1, \cdot) = Q(j_2, \cdot)$. Given such a kernel Q, ${}^{(\gamma)}Q = Q(\gamma \cdot \gamma^{-1}, \gamma \cdot \gamma^{-1})$ is clearly also a kernel on J; it remains to show that if $\pi(j_1) = \pi(j_2)$ then $Q(\gamma j_1 \gamma^{-1}, \cdot) = Q(\gamma j_2 \gamma^{-1}, \cdot)$. But to say that π is a bundle for L acting on J by conjugation means that $\pi(j_1) = \pi(j_2) \Rightarrow \pi(\gamma j_1 \gamma^{-1}) = \pi(\gamma j_2 \gamma^{-1})$, so the desired result follows from the property of Q.

Now, to prove 5.12, we begin with the kernel ${}^{\gamma}\langle Q_1, ..., Q_k\rangle_\tau$ on E^k. For $e = (e_1, .., e_k) \in E^k, \Delta = \Delta_1 \times ... \times \Delta_k \in \mathcal{E}^k$,

$${}^{\gamma}\langle Q_1, ..., Q_k\rangle_\tau(e, \Delta) = \langle Q_1, ..., Q_k\rangle_\tau(\; \gamma^{-1}e, \; \gamma^{-1}\Delta)$$

by 5.1 and **7**–4.1. This last expression is

$$\sum_{\chi \in \mathcal{I}(k)} \tau(\gamma^{-1}e_1, ..., \gamma^{-1}e_k; \chi)$$
$$\prod_{i \in D(\chi)} Q_i\big((\gamma^{-1}e_{\chi(i)})(\gamma^{-1}e_i)^{-1}, (\gamma^{-1}\Delta_i)(\gamma^{-1}e_i)^{-1}\big) \prod_{i \notin D(\chi)} \epsilon_{\gamma^{-1}e_i}(\gamma^{-1}\Delta_i)$$

$$= \sum_{\chi \in \mathcal{I}(k)} \tau(e_1, ..., e_k; \chi)$$
$$\prod_{i \in D(\chi)} Q_i\big((\gamma^{-1}e_{\chi(i)})(\gamma^{-1}e_i)^{-1}, (\gamma^{-1}\Delta_i)(\gamma^{-1}e_i)^{-1}\big) \prod_{i \notin D(\chi)} \epsilon_{\gamma^{-1}e_i}(\gamma^{-1}\Delta_i)$$

since τ is translation invariant. Recall that $(\gamma^{-1}e_{\chi(i)})(\gamma^{-1}e_i)^{-1}$ denotes that element $j \in J$ such that $j(\gamma^{-1}e_i) = \gamma^{-1}e_{\chi(i)}$. It is then evident that

$$(\gamma^{-1}e_{\chi(i)})(\gamma^{-1}e_i)^{-1} = \gamma^{-1}(e_{\chi(i)}e_i^{-1})\gamma,$$

and similarly

$$(\gamma^{-1}\Delta_i)(\gamma^{-1}e_i)^{-1} = \gamma^{-1}(\Delta_i e_i^{-1})\gamma.$$

Moreover $\gamma^{-1}e_i \in \gamma^{-1}\Delta_i \iff e_i \in \Delta_i$, so that

$$\epsilon_{\gamma^{-1}e_i}(\gamma^{-1}\Delta_i) = \epsilon_{e_i}(\Delta_i).$$

Thus, the last expression above may be written

$$\sum_{\chi \in \mathcal{I}(k)} \tau(e_1, ..., e_k; \chi) \prod_{i \in D(\chi)} Q_i\big(\gamma^{-1}(e_{\chi(i)}e_i^{-1})\gamma, \gamma^{-1}(\Delta_i e_i^{-1})\gamma\big) \prod_{i \notin D(\chi)} \epsilon_{e_i}(\Delta_i)$$

$$= \sum_{\chi \in \mathcal{I}(k)} \tau(e_1, ..., e_k; \chi) \prod_{i \in D(\chi)} {}^{(\gamma)}Q_i((e_{\chi(i)}e_i^{-1}), \Delta_i e_i^{-1}) \prod_{i \notin D(\chi)} \epsilon_{e_i}(\Delta_i)$$

$$= \langle\, {}^{(\gamma)}Q_1, ..., {}^{(\gamma)}Q_k\rangle_\tau(e, \Delta). \quad \blacksquare$$

Scholium 5.13. Let a group Γ act measurably on the left on a space Z. In 5.1, for $\gamma \in \Gamma$ we considered the linear operation on the function space $\mathcal{Z}$ induced by $z \to \gamma z$; we used the same symbol γ to denote this linear operator: $(\gamma f)(z) = f(\gamma^{-1}(z))$. Thus there is an induced left action of Γ on functions, namely $f \to \gamma f$ (or ${}^\gamma f$). Now suppose Γ acts on Z on the right. We can consider the linear operator on functions induced by $z \to z\gamma$, and we also get a left action on functions, namely $f \to \gamma f$, where now $(\gamma f)(z) = f(z\gamma)$. If Γ acts on Z both on the left and right, then we will use the notation $(\gamma_l f)(z) = f(\gamma^{-1}z)$. and $(\gamma_r f)(z) = f(z\gamma)$. For example we consider our group J' acting on itself by multiplication on both the left and right. γ_l and γ_r are distinct in general (unless J' is abelian), but they commute. If we view kernels $Q: J \times \mathcal{J} \to [0, 1]$ as operators on functions in the usual way, then we can express the left action $Q \to {}^{(\gamma)}Q$ as follows:

$${}^{(\gamma)}Q = (\gamma_l\gamma_r)Q(\gamma_l\gamma_r)^{-1}.$$

Example 5.14. Suppose that in the framework Θ we have $S = J/H$ for a subgroup H of J, and $\pi: J \to S$ is the canonical projection; these are frameworks like those of Example **5**–4.1. Let $L \subset J$ be any subgroup contained in the normalizer of H in J, i.e., L is any subgroup of J in which H is normal. Then $\pi: J \to S$ is also a bundle for the action of L by conjugation. In fact, a fibre of π is a coset jH, and for $l \in L$ we have $l(jH)l^{-1} = (ljl^{-1})H$ (since $Hl^{-1} = l^{-1}H$) which is another coset. Thus conjugation by l permutes the fibres of π as claimed.

Suppose that we have a framework Θ with a τ-distribution which satisfies the hypotheses of Proposition 5.10. We would like to construct a framework Θ' which is a specialization of Θ, in which (with notation as in 5.8 through 5.11) E' and S' are the distinguished configuration and premise spaces, and J' is the distinguished symmetry group. We assume that the action of J' on the set E' of participator ensembles is compatible with the action of J' on action kernels given by 5.11. More explicitly, we can start with an "initial" participator ensemble

$$A = \{A_i\}_{i=1}^{k}, \quad A_i = (\xi_i,\ Q_i, \eta_i).$$

Then we assume that

$$^{\gamma}A = \{\ ^{\gamma}A_i\}_{i=1}^{k}, \quad ^{(\gamma)}A_i = \left(\ ^{(\gamma)}\xi_i,\ ^{(\gamma)}Q_i,\ ^{(\gamma)}\eta_i\right) \tag{5.15}$$

where $^{(\gamma)}Q_i$ is as in 5.11. The action of γ on the ξ_i and the η_i is assumed given, but we need not stipulate its properties now.

Now to build Θ', let us assume that we have chosen some set X' of (say k-fold) participator ensembles which contains E', and some set Y' of measures on E^k which contains S'. We further assume that we have a group G' which contains J' and acts on X' in a manner which extends the action of J' on E'. For simplicity, however, we will focus our attention only on the distinguished part of the structure, E', S', J'. We can define a fundamental map $\pi': J' \to S'$ using $\pi': E' \to S'$ and our "initial" element $A \in E'$:

$$\pi': J' \to S',\ \pi'(\gamma) = \pi_0'(\ ^{(\gamma)}A) =\ ^{(\gamma)}\nu_0. \tag{5.16}$$

In this way we get a symmetric framework

$$\Theta' = (X', Y', E', S', G', J', \pi'). \tag{5.17}$$

We now discuss the question: Is Θ' is a specialization of Θ in the sense of 3.2? The primary issue here is the nature of the specialization scheme 3.1. Our distinguished configurations E' are already described as a set of participator ensembles on Θ, namely the set of all the ${}^{(\gamma)}A$. De facto, in any specialization scheme which applies to this situation, participator ensembles of this type must satisfy the permissibility condition of the scheme. Recall that the role of the permissibility condition is to ensure two things: first that the separate permissible participator ensembles have asymptotically stable dynamics; second that the perturbation of these stable asymptotic characteristics of one such ensemble by its interaction with another has sufficient regularity to encode information about the interaction in accessible form. Then, according to Definition 3.2, the possible regular perturbations which arise in this manner must be parametrized by S', and they must encode accessible information about the interactions in a very precise sense: Suppose the two ensembles ${}^{(\gamma_1)}A$, ${}^{(\gamma_2)}A$ correspond to points e_1', e_2' of E'. The interaction of these ensembles perturbs the asymptotics of ${}^{(\gamma_1)}A$ in a manner which is encoded by the element $\pi'_{e_1'}(e_2') = \pi'(\gamma_2\gamma_1^{-1})$.

We want to see what this means in our case. Using the same notation as in 5.4, let us denote the transition probability of the initial participator ensemble A by P_0, so that

$$P_0 = \langle Q_1, ..., Q_k\rangle_\tau.$$

We have fixed a stationary measure for P_0 on E^k, denoted ν_0, and

$$S' = \{\ {}^{(\gamma)}\nu_0 \mid \ \gamma \in J'\}.$$

According to the definition of π' given in 5.16, $\pi'(\gamma) = {}^{(\gamma)}\nu_0$. Thus the perturbation regularity requirement may be stated as follows.

5.18. The interaction of the ensembles ${}^{(\gamma_1)}A$ and ${}^{(\gamma_2)}A$ perturbs the asymptotics of ${}^{(\gamma_1)}A$ in a manner which is encoded by the measure ${}^{\gamma_2\gamma_1^{-1}}\nu_0$.

Broadly speaking, there are two ways in which 5.18 might hold: concretely, and abstractly. In the concrete way the perturbation information is encoded in the properties of the measure ${}^{\gamma_2\gamma_1^{-1}}\nu_0$ *as a measure*. In the abstract way the information is simply encoded in the group element $\gamma_2\gamma_1^{-1}$ which is attached to the measure. How might the concrete way work? Recall that for every $\gamma \in J'$, ${}^{\gamma}\nu_0$ is stationary for ${}^{\gamma}P_0$; by 5.2 and 5.12 this kernel is the transition probability for the ensemble ${}^{(\gamma)}A$. Under suitable hypotheses (say on the initial ensemble A) the interaction in question may perturb ${}^{(\gamma_1)}A$ so that its stationary measure ${}^{\gamma_1}\nu_0$ is canonically deformed toward ${}^{\gamma_2}\nu_0$, and moreover so that the measure ${}^{\gamma_2\gamma_1^{-1}}\nu_0$ is some type of derivative of this deformation. This possibility is mathematically appealing. On the other hand, the abstract way for 5.18 to hold requires only that *some* canonically specified aspect of the total perturbation may be classified by the group-theoretic difference $\gamma_2\gamma_1^{-1}$ between the two interacting systems. We do not analyze these questions further here, and in fact no definitive analysis is now available. In the next chapter, when necessary, we will simply assume 5.18 is satisfied, so that we have a bona fide specialization scheme for which Θ' is a specialization of Θ.