Solutions to the Binding Problem: Progress through Controversy and Convergence

Anne Treisman*
Psychology Department
Princeton University
Princeton, New Jersey 08544

The binding problem, or constellation of problems, concerns our capacity to integrate information across time, space, attributes, and ideas. The goal of research in this area is to understand how we can respond to relations within relevant subsets of the world but not to relations between arbitrarily selected parts or properties. Language comprehension and thinking critically depend on correct binding of syntactic and semantic structures. Binding is required when we select an action to perform in a particular context. We must, for example, reach in the right direction, lift the glass with the correct muscle tension, and drink the water it contains rather than eat or inhale it. The mediating “event file” binds stimulus to response (Hommel, 1998; Shadlen and Movshon, 1999 [this issue of Neuron]). The most extensive discussion so far has focused on the problem of binding in visual perception. How does the brain segregate the correct sensory data to represent the objects that are actually present and not some illusory recombinations of their features? In considering perceptual binding, it is important to note that “seeing” an object is not the same as identifying it (Kahneman et al., 1992; Treisman, 1992a; Treisman and Kanwisher, 1998). To generate a perceptual experience and to be able to act on it, we need to specify the current details of how an object looks, where it is, how it is oriented, and many other often arbitrary details of its current instantiation. Thus we must construct a temporary token (“object file”) that binds together these current features with the more permanent identifying characteristics of its type.

The present papers have dealt mostly with this perceptual version of the binding problem, although many ideas can apply more broadly to binding at all levels. Why should a binding problem arise in vision, both for the brain and for the scientist attempting to understand it? A number of factors could contribute to create binding failures in vision. One is that various properties of objects appear to be separately analyzed by specialized subsystems. Thus, while information from the same location is implicitly bound by the cells that respond to it initially (allowing the selective adaptation to conjunctions of features described by McCollough, 1965; see also Wolfe and Cave, 1999 [this issue of Neuron]), at later stages the information from these cells appears to be routed to different neural populations, forming a distributed representation of an object’s different properties. Another is that receptive fields at higher levels are large enough to generalize across a wide range of locations. Because visual scenes typically contain multiple objects, the question of which features belong to which objects could frequently arise. To aggravate the problem, different parts of a single object occupy different locations, and there may be occluding objects that break their continuity. Coarse coding of different stimulus dimensions also creates representations that may depend on ratios of activity in neurons with different but overlapping tuning. Whenever the perceptual representations of simultaneously present objects depend on distributed patterns of firing in populations of cells, the risk of superposition ambiguities within the same neural network will arise, creating a need to identify and to signal which units belong to the same representation.

For any case of binding, the binding problem can actually be dissected into three separable problems. Different theories have focused primarily on one of the three.

1. Parsing. How are the relevant elements to bind as a single entity selected and segregated from those belonging to other objects, ideas, or events?
2. Encoding. How is the binding encoded so that it can be signaled to other brain systems and used?
3. Structural description. How are the correct relations specified between the bound elements within a single object?

The second and third operations are not necessarily sequential, and in fact some models combine all three as part of the same process. I will discuss each of these three aspects of the binding problem in turn.

Parsing
There are a number of different ways in which the initial parsing of objects might be accomplished. In particular, the selection of parts and of properties may depend on different mechanisms. Different attributes of the same object, such as its color, orientation, and direction of motion, must occupy the same location. However, different parts, like the arms and legs of a child or the two colors of her shirt and pants, occupy different locations and, if she is partly occluded, may not even be spatially linked in the retinal image. Possible ways of parsing objects and backgrounds include the selection of sensory data that share the same temporal parameters (onset, offset, flicker rate), or that match a prestored template (e.g., Chelazzi et al., 1993), or that occupy the same or adjacent locations, or finally that share one or more Gestalt properties, such as their color or texture, common fate, collinearity, good continuation, symmetry, and convexity.

In cases where objects have different temporal onset or offset times, the externally imposed synchrony of the initial neural response to one object may solve the parsing problem as well as the encoding problem (see below). The evidence suggests that temporal modulations can be used to separate figure and ground (Leondards et al., 1996). However, when pitted against spatial cues temporal modulations do not disrupt selection by location (Fahle and Koch, 1995; Kiper et al., 1996), so they cannot always be the dominant cue.

When one of the objects is known or expected, the

*E-mail: treisman@princeton.edu.
selection may be mediated by a match to this familiar or cued object. The best-known example is the Dalmatian dog to be extracted from a background of black and white patches (Figure 1). One of two superimposed movies of real life scenes (like a ball game and a hand game) can be selected so efficiently that a salient event in the unattended movie (a lady with an umbrella walking across the field) is completely missed (Neisser and Becklen, 1975). This type of segregation no doubt uses other grouping cues such as common fate, collinearity, and shared colors, but top-down predictions are also likely to play a role. Zhang (1999) showed recently that suppression of irrelevant stimuli can also be helped by precueing the template of the unwanted stimulus, even when it occupies the same location as the target, ensuring that suppression is not mediated by location but by a representation of the object itself.

Most of the research on binding has been devoted to the last two forms of selection—by location and by shared features. Feature Integration Theory (FIT; e.g., Treisman and Gelade, 1980; Treisman, 1993; Treisman, 1998) uses both. It was developed to account for a number of empirical findings: (1) that search for targets that need binding to distinguish them from the nontargets often requires attention; (2) that when attention is directed elsewhere, illusory conjunctions wrongly recombining features of different objects are frequently seen; (3) that precueing the relevant location helps much more when a conjunction must be reported than when the targets are defined as a conjunction of separate features; and (4) that grouping by single features occurs in parallel across the field, whereas grouping by conjunctions is much less salient and also seems to require attention. In early papers, we proposed that binding is achieved by directing spatial attention serially to the locations of different objects (or homogeneous groups of objects). Features of objects in unattended locations are thereby excluded and cannot form illusory conjunctions with the features of the attended object. The relevant locations are selected in a “master map” of locations by an externally directed “window of attention” serially focused on single filled locations or contiguous clusters that might correspond to “objects.” The attention window gives access to the features in the corresponding locations in the different feature maps and allows the information from those locations to be assembled in a single object file for further analysis and identification.

The separation of explicit access to features and to locations may correspond to the separation of ventral “what” and dorsal “where” pathways, although there must be implicit links between the two as well as implicit location information in the feature maps. The fact that patients with bilateral parietal lesions (Balints’ syndrome) have major problems with binding is consistent with the idea that the master map of locations is associated with parietal function (Robertson et al., 1997). If our model of the deficit is correct, the simultanagnosia that these patients also suffer is evidence for the importance of binding to normal object perception. If only one object can be bound, it seems that only one object is seen.

Treisman (1988; Treisman and Sat0, 1990) added a feature-based selection process to the original version of FIT to account for cases where highly discriminable features appear to group and allow rapid or parallel conjunction search. The idea was that the binding process could be bypassed if connections from the separate feature maps responding to the target features were used to signal the corresponding locations within the master map and to suppress all others, thus excluding all nontargets from further processing. For example, if a target of a search is known to be red and vertical among green vertical and red horizontal lines, all the active locations in a color map for red and an orientation map for vertical could select the corresponding locations in the master map through the implicit links connecting them, and temporarily inhibit all other locations. Wolfe et al. (1989) proposed a similar model of Guided Search, and Wolfe and Cave (1999) report evidence supporting the feature-based account. This form of selection may also be responsible for grouping the parts of partially occluded objects.

In a different form of grouping by Gestalt properties,
edges or elements that are continuous or collinear appear to be linked by horizontal connections in area V1 (Kapadia et al., 1995), which may increase the salience of object boundaries (Yen and Finkel, 1998). These connections might also be used to define the locations to be selected or suppressed in the master map of locations. Constraints on the shape of the attention window could also bias the selection in favor of simple, symmetrical, convex objects.

Several papers in this issue develop similar ideas to explain selection by location or feature grouping. Wolfe and Cave present their Guided Search model (Wolfe et al., 1989; Cave and Wolfe, 1990; Wolfe and Bennett, 1997), which extended FIT. They recently modified the earliest stages to include an initial loose “bundling” by location, as opposed to the “tight binding” achieved through attention. Shadlen and Movshon favor the related computational model proposed by Olshausen et al. (1993), using shifter circuits to achieve the same goal of serial processing by locations. Reynolds and Desimone adopt the idea that spatial attention is the main mechanism for binding, combining it with their biased competition account and linking it to evidence from single unit-recordings (see also Luck et al., 1997b).

Reynolds and Desimone (1999 [this issue of Neuron]) also propose salience and grouping as supplementary binding mechanisms that could, by the increased activation they cause, select a winner in a competitive interaction with other objects. One concern with this hypothesis is that, without additional specification of the selection process, it seems to require the binding problem to be at least partly solved in order to contribute to its solution. The attribute on which the stimuli are salient (e.g., luminance or color contrast), or on which they are grouped (e.g., color, orientation, shared motion), must be bound to the other attributes of the same stimuli before their greater activation can be transmitted to those other dimensions. This “catch 22” might be resolved in cases where salience and grouping are determined at the early levels of processing, before the stimulus attributes are segregated into separate specialized visual areas. At these earliest stages, high contrast may ensure high firing rates for holistic stimuli rather than for particular attributes, so that subsequently each attribute wins out in the competition within its own specialized area. The mechanism is less clear for cognitively determined or learned salience and grouping (such as the salience of one’s own name in a list of other names). The identity of the name might not be available at stages that precede the specialized processing of separate attributes.

In order to make progress in understanding, it is useful to sharpen the disagreements so that clear distinctions in empirical data can decide between them. As I understand it, Reynolds and Desimone’s biased competition model differs from FIT primarily in the source of suppression of unwanted stimuli, ideas, or responses. (The main difference claimed by Reynolds and Desimone was that illusory conjunctions, in their model but not mine, arise from spatial uncertainty within receptive fields, but this was in fact the model proposed by Treisman and Gormican, 1988 [see pp. 45–46].) I also proposed additional mechanisms besides spatial attention that can be used to select stimuli and deal with the binding problem, one being grouping and one being object tracking by reentrant connections [Treisman, 1995]. Both accounts assume an external source of control by selective attention, presumably directed by prefrontal and parietal areas. But Reynolds and Desimone restrict this control to biasing the competition that would occur between objects anyway. They assume that selection of one stimulus or response from many is directly determined by suppressive links between their neural substrates, with the more active winning over the less active. Attention in their model biases the competition by adding top-down activation to one of the competing sets of cells.

FIT, on the other hand, attributes an inhibitory as well as an activating role to the external control system that we label attention. The evidence Reynolds and Desimone cite is the observation that when a second stimulus is introduced within the receptive field of a visual neuron, its response is a weighted average rather than the sum of the two separate responses (see also Miller et al., 1993). This is consistent with a direct suppressive interaction between the two afferent inputs to the cell, but it could also result from external selection of one and inhibition of the other to reduce cross-talk and binding errors. Intrinsic competition may be difficult to distinguish from the effects of dividing or focusing an extrinsic source of attention. One piece of evidence that favors the biased competition model is the finding by Reynolds et al. (1999) that when two unattended stimuli are presented in the same receptive field, the response is still lower than the response to the more effective of the two presented alone, even though they are not competing with each other for attention. Of course, this does not preclude the idea that attention also inhibits unwanted stimuli when binding errors might otherwise occur.

If the idea of intrinsic competition is correct, it raises many interesting questions for research. How do the cells “know” that they are being activated by different objects rather than by one complex object? There may be some feedback from higher object recognition areas, but when both objects are unattended, this is likely to be limited. Does similarity play a role in determining the degree of competition? What other factors affect the competition and its outcome?

Encoding

Several of the present papers deal primarily with the second problem: how the feature bundles, once selected, are encoded to be used for thinking, deciding, and acting. Essentially, this is a special case of the central question, “What, in neural terms, corresponds to the final representation of what we see?” Is it the activation of particular labeled cells, or particular cell assemblies, or particular temporal patterns of activity within or across cells, independent of which cells implement the pattern? Or is it a combination of place and temporal pattern? We are far from having an answer, which is quite a handicap in devising models of object perception, but the question has not been discussed much. One constraint, of course, is that the codes should remain distinct when several objects are present at once (the binding problem).

The focal hypothesis debated in the present issue is the proposal by Milner (1974) and von der Malsburg...
(1981) that the neurons coding elements that belong to the same object are distinguished from others by firing in synchrony. Oscillations in the range of 30–60 Hz are thought to assist and perpetuate the synchronization, especially for widely separated neurons. While the synchrony lasts, the cells that share it are treated as representing the same object, event, or proposition. In this issue, Von der Malsburg (1999) gives the theoretical reasoning behind models of binding by synchronized firing. Gray (1999) and Singer (1999b) discuss the physiological evidence and implementation of these ideas. In contrast, Shadlen and Movshon (1999) and Ghose and Maunsell (1999) point out some problems for the synchrony account, including the supposedly limited precision of temporal and spatial coding by neurons, the need to use timing relations to represent real temporal differences (in such discriminations as the perceived direction of sounds or the syllables of speech), and the failure in several studies to observe the predicted oscillations in visual areas (e.g., Tovee and Rolls, 1992a; Young et al., 1992).

The binding-by-synchrony hypothesis has created considerable interest and excitement, since it provides a means of disambiguating superimposed distributed codes in neural networks, thus greatly increasing their flexibility. It also provides a plausible reason for the attentional limit of around four objects that is widely observed in the perception of brief displays and in studies of visual working memory: the different firing rates that can be easily discriminated on a background of inherent noise and accidental synchronies may set a low limit to the number of objects that can be simultaneously bound.

Von der Malsburg (1999; see also Singer, 1999b) points out another important advantage to temporal binding which is not often discussed: it allows coarse coding within dimensions. Coding intermediate values on perceptual dimensions by ratios of activity in differently tuned but overlapping populations of cells can maximize both neural economy and discriminability. However, if different values are signaled by particular combinations of cells, the binding problem reemerges as soon as more than one coarsely coded feature is present. While exploring the effects of coarse coding and of similarity on binding, we had found that the predicted illusory conjunctions can indeed arise within dimensions (e.g., illusory purple with brief presentations of red and blue), and that attention seems to play the same role for within-dimension binding as for between-dimension binding (Treisman, 1991, 1992b).

The main alternative hypothesis for signaling the outputs of the binding process is a place code, which represents different objects or parts of objects by the firing of different labeled conjunction coding or “cardinal” cells at the top of a hierarchical perceptual system (Barlow, 1972, 1985). The cardinal cells could be replaced by cell assemblies, provided that the coding is sufficiently sparse for overlap between cells taking part in different assemblies not to be a problem. The place or “labeled line” hypothesis is developed in this issue by Shadlen and Movshon, by Ghose and Maunsell, and by Riesenhuber and Poggio, who created a model to demonstrate its feasibility.

There is considerable evidence for coding of specific percepts by specialized cells. Single units in monkeys respond to faces better than to other stimuli (e.g., Perrett et al., 1985). The behavioral discriminations of motion made by a monkey can be predicted from the activity of individual cells in area MT (Newsome et al., 1989; Britten et al., 1996)—quite strong evidence for cardinal cells in the case of directions of motion. However, faces may be a special case, highly significant from the evolutionary point of view, and objects are typically more complex than directions of motion. Ghose and Maunsell suggest that the number of objects we can actually identify is only in the tens of thousands, but we can also see thousands of differences between individual tokens and views of these objects. The cardinal cell hypothesis does run into combinatorial explosion problems if all discriminable instantiations of all objects must have unique cells to signal their presence.

A related difficulty is that this account allows no distinction between identifying and “seeing,” or between types and tokens (Kahneman et al., 1992; Treisman, 1992a). Using cardinal cells, feature binding can be coded only by identifying pre-stored conjunctions. It is not clear, therefore, how a new and unexpected conjunction would be bound and perceived. Consider seeing a three-legged camel with wings, or a triangular book with a hole through it, or a new object like an electron microscope picture of a cell that we have never seen before. All these would be instantly visible and bound, even when we have no idea what they are. The plasticity of the nervous system may allow new cardinal cells to be created through learning and experience—for example, to signal familiar configurations like letters, digits, and grandmothers. But novel objects can clearly be “seen” under good conditions the first time they appear.

Another deficiency is that cardinal cells have no way to represent the hierarchical structure of the object, the different momentary constellations of articulated parts in a given token as seen on a particular occasion. These could perhaps be read from the complete pathway through the hierarchy that culminates in the cardinal cell. Ghose and Maunsell (1999) suggest that “neurons throughout visual cortex could contribute equally, so that neurons in earlier stages could be involved in discrimination of low-level attributes such as precise positions or orientations, and those in later stages involved in the assessment of complex patterns and forms.” But if this option is adopted, the binding problem may re-appear. How would the parts of different objects be kept distinct if both are active at once? One would need a way of storing the sequence of synapses that activated a cell, and where its afferent connections came from.

It is fascinating to see extremely strong cases made both for and against the two theories. It allows readers to weigh the evidence presented and reach their own informed conclusions. The main controversy concerns the plausibility of the synchrony hypothesis. So far, there is empirical evidence supporting both sides, and in some cases the conclusions are conflicting. For example, in binocular rivalry, perceptual dominance in awareness has been shown to correlate on the one hand with which neurons are firing (Logothetis, 1998; Tong et al., 1998) and on the other with synchronized firing (Fries et al., 1997; Tononi and Edelman, 1998). Some studies have
shown impaired perception when externally imposed oscillations cross figure boundaries (Leonards and Singer, 1998), but others have not (Fahle and Koch, 1995; Kiper et al., 1996). Some have shown advantages when externally imposed synchrony coincides with object boundaries, and others have not (Blake and Tang, 1997). Some have found the predicted synchrony for receptive fields within figures but not across figure and ground, and others have not (Lamme and Spekreijse, 1998). There are physiological arguments and evidence that seem convincing on both sides. The fact that the two mechanisms—synchronized circuits and specialized conjunction cells, at least for object components (Tanaka, 1996)—are compatible allows a strong possibility that both play a part in binding.

One of the benefits of discussions such as these is that they suggest new experiments that could provide important data. Shadlen and Movshon have an interesting proposal to test whether the output of the binding process really is synchronized firing. This is to record from the same pairs of single units when different figures are in their receptive fields and when the same bent figure falls on both receptive fields, and to do this both when the linking area is occluded and when it is not. Singer describes an as yet unpublished study, in which cells responding to orthogonal directions of motion synchronized their activity when the stimuli were consistent with the motion of a single plaid surface and not when the stimulus was seen as separate components moving in two different directions (Castelo-Branco et al., 1998, Eur. J. Neurosci., abstract). Gray proposes repeating an experiment like that reported by Lamme and Spekreijse (1998), recording from receptive fields either both on a Gestalt-induced figure or one on and one off the figure. The additional manipulation that he proposes is that the degree of perceptual segregation shown in the animal’s behavior on the same trial as the recording should be correlated with the degree of synchronization observed. Another possibility that occurs to me would be to present illusory Kanizsa figures induced by pacmen (e.g., Figure 2) and to record from two receptive fields on different illusory edges of the same figure. The prediction would be that their synchrony would be reduced or abolished when the pacmen are rotated to create two new and separate figures (with other pacmen) and to destroy the shared illusory surface that united them. If synchrony is observed with the shared figure, these experiments would have the advantage that the stimuli within the receptive field would be identical in the two conditions.

Structural Description
The specification of how the elements, once selected, should be bound in the correct structured relations, is dealt with extensively by a number of psychological theories: theories of object identification and recognition (e.g., Biederman, 1987; Bülthoff et al., 1995; Tarr and Bülthoff, 1995) deal with structural relations within objects, linguistic theories explain structural relations within sentences and propositions, and theories of perceptual-motor control specify the structural relations between stimuli and actions.

Wolfe and Cave (1999) describe some experiments which suggest that binding problems can arise in forming the structural description—more specifically, in determining how parts within an object should be bound and how illusory conjunctions within objects can be avoided, preventing errors such as an exchange of locations between the eyes and the mouth in a face, or between the colors of a child’s shirt and pants. Several of the solutions described in the Parsing section, such as suppressing other items in the display, or synchronizing the firing rates of cells responding to the eyes and the mouth, would not solve the within-object binding problem. One answer may be to attend to each part within the object in turn, in order to bind them to their within-object locations. Favoring this view is the fact that search latencies involving within-object discriminations of this type, for example between Ts and Ls in random orientations, increase linearly with the number of items in the display. Perhaps the reason why these searches are typically slower than those for between-object conjunctions is not only that feature guidance cannot be used, but also that attention must be more narrowly and precisely focused within the object than between objects.

Some recent findings by Tallon-Baudry and Bertrand (1999) in human participants suggest that the oscillations that can be observed in electroencephalogram (EEG) records may be associated with binding in the formation of object representations. They found an increase in the oscillation range both when the Dalmatian dog emerged from the background of spots in the well-known picture (Figure 1) and when a Kanizsa triangle was seen occluding three black disks (Figure 2), but not when the disks were rotated to abolish the figure. In both cases, stimulus elements (a subset of dots, the three pacmen) must be selected and bound before the figure can be seen and its three-dimensional structure inferred. The observation of EEG oscillations in association with controlled perceptual tasks offers a promising new tool to study binding in humans. Timing their onset, for example, may help determine at what stage of processing they arise. It would also be useful to see if separate frequencies can be simultaneously detected, perhaps in different hemispheres, when two objects are presented in the left and right visual field.

Models using cardinal cells to encode binding often encode the structural relations of parts as part of the
process of reaching the cardinal cells. An example is the model described in this issue by Riesenhuber and Poggio (1999a), who claim it as an existence proof that binding poses no real problem for vision. The model is based on the hierarchical architecture described by Hubel and Wiesel (1968). The novel feature of Riesenhuber and Poggio’s proposal is that, in addition to the traditional conjunction coding by convergence, a generalization pooling of similar features across positions and views should use a MAX operation that picks the most active afferent unit as the only one to control the response at any point. Riesenhuber and Poggio show that this reduces interference from extraneous activation as well as solving the generalization problem across places, views, sizes, and so on. The binding problem is solved by first coding very local relations through the combination of cells with adjacent small receptive fields at early stages, and then progressively generalizing across increasing areas for successively more complex encoded conjunctions.

Riesenhuber and Poggio’s model is certainly interesting and solves some binding problems, but I think it may still run into some of the same difficulties as other place coding accounts. It is not clear that it really solves the combinatorial explosion problem. So far, it has been tested only on restricted sets of objects. I am also not sure that behavioral evidence supports the local-to-global sequential ordering, either for the relative speed of responses (Navon, 1975) or in the determinate order of processing. Oliva and Schyns (1997) showed that processing order could be strategically determined by the task or by priming high or low spatial frequencies. The experiments described by Wolfe and Cave also suggest that Riesenhuber and Poggio’s model may actually do better than humans at solving within-object binding problems. Finally, the model does not distinguish identifying from seeing, thus raising the problems mentioned in the Encoding section. However, in combination with synchronized firing of cells activated from the same location or from shared features, this model might provide a more complete solution to the problems of perceptual representation.

Further Questions
Obviously, the central question for the future is which of the proposed mechanisms at any of the three stages are the most likely to be correct. Note that several of the papers in this issue postulate more than one binding mechanism. It seems very likely that all play a part, but in which circumstances and which combinations remains to be determined.

Another question concerns how models for parsing, encoding, and structural description are best combined to form a coherent integrated account of binding. It may be possible to combine any hypothesis for one operation with any of the hypotheses for the others. For example, the choice of which sense data to encode by synchronous firing could be mediated by spatial attention. So far, there has been little attempt to discuss the relative plausibility of the different possible combinations. One example of an integrated model, combining spatial selection and grouping with synchrony, was proposed by Hummel and Biederman (1992) and elaborated by Hummel and Stankiewicz (1996). Riesenhuber and Poggio’s model makes selection, encoding, and structural description all part of the same single process of place coding through a local-to-global hierarchy.

Perhaps the most crucial question is this: how is the output of the binding process interpreted and used by later stages? One constraint is that the outputs must be readable by whatever systems encode memories, determine choices, or select motor responses. Moreover, the link to responses must be highly flexible. We have the ability to arbitrarily choose to attach most possible responses to most perceptual codes and to rapidly change the pairing, according to circumstances, instructions, or personal whim. Probably the simplest solution from this point of view would be the “labeled line” idea, where different percepts correspond to the firing of different cells. But this runs into the other problems discussed above. If synchronized firing is hypothesized, the question of how the synchrony is read by later stages also remains to be answered. Singer points out that synchronous firing gives cells an advantage in activating cells at subsequent stages of processing. Von der Malsburg suggests that coincidence detectors in later circuits could resonate to particular correlated patterns of firing, and that short term reversible changes in synaptic efficiency could be used temporarily to eliminate connections between uncorrelated units. Further research will be needed to explore in more detail how our brains can so flexibly bridge the gap between percept and action.

References
A comprehensive reference list for all reviews can be found on pages 111–125.
References for Reviews on the Binding Problem


