Most prominent models of visual object recognition (e.g., Biederman, 1987; Marr, 1982) view the apprehension of a particular object in an image as exclusively based on a data-driven preconceptual recovery of the object’s structural features (e.g., geons, contour-segments, etc.) from the image. These accounts of an object-limited mapping between pattern and object meet most central demands plausible models of human object perception should meet. They offer speed, resistance to noise in the image and differences in object orientation, applicability to a wide variety of objects without losing discriminatory power, and so on. However, they do neglect one typical characteristic of everyday human object perception: objects are virtually always encountered in the context of a scene.

Often this has been taken to merely pose an additional problem of image segmentation and body-finding that needs to be solved prior to the actual process of object recognition (e.g., Waltz, 1975), which then is assumed to be identical for all objects whether appearing in isolation or in a full scene. More recently, however, evidence is accumulating in support of the position that for objects in natural, real-world scenes the pattern of light reflected by the object itself may not be the sole source of object-diagnostic information used in human object recognition.

In fact, a number of studies by Biederman and colleagues (Biederman, 1981; Biederman, Mezzanotte, & Rabinowitz, 1982; Biederman, Mezzanotte, Rabinowitz, Francolini, & Plude, 1981; Klatsky, Teitelbaum, Mezzanotte, & Biederman, 1981) appear to indicate that this pattern information may only be of secondary importance. In Biederman’s studies, subjects are asked to verify whether a prenamed object was present at a cued position in a very briefly presented scene. Speed and accuracy of responses in this speeded target verification task are assumed to directly reflect the perceptibility of the object at the cued position. Guided by this rationale, results obtained with the speeded target verification paradigm have fostered the claim that an object’s perceptibility is strongly affected by the degree to which its spatial and semantic relations to its context conform to what they typically are in the real world. Specifically, it was
found that target verification was slower and less accurate when the cued object (a) inappropriately floated, (b) appeared in an unlikely scene, (c) had an inappropriate size, or (d) was in an inappropriate position. These effects of Support, Probability, Size, and Position violations were observed for cued objects both in foveal and extrafoveal vision. Consequently, there seems to be good reason to defend the position that contextual information is not only a mandatory and integral part of the effective stimulus (Tiberghien, 1986) for object recognition, but also is sufficiently diagnostic to reliably steer this process even in the absence of detailed object pattern information. Recently, this position has been reaffirmed by Boyce, Pollatsek, and Rayner (1989). Using the same paradigm, they found that verification of extrafoveally presented objects was affected by the semantic consistency between the objects and the background of the scene they appeared in.

In combination with reports of a chronological precedence of global scene characterization over individual object recognition (e.g., Antes, Penland, & Metzger, 1981; Loftus, Nelson, & Kallman, 1983; Metzger & Antes, 1983), these findings appear to have two important implications for the currently dominant data-driven models of visual object recognition. First, they raise the question of whether object recognition under everyday circumstances can indeed be viewed as a completely modular process, limited to a structural analysis of the pattern reflected by the object. Second, they caution against the assumption implicit in much vision research, that bottom-up models of individual object recognition constitute the core of more general models of scene perception. In Figure 14.1, I have attempted to outline this object-based view on real-world perception using the computational framework proposed by Marr (1982).

As can be seen in this figure, a scene only becomes meaningful by virtue of the identification of the various objects in it, and the semantic interpretation of the spatial relations that hold between these objects. In other words, this view holds that scene comprehension is the result of the construction of an ad hoc “situation model” (cf. Van Dijk & Kintsch, 1983) which interfaces the specific scene that is being viewed with general world knowledge. The research mentioned above, however, suggests that more holistic scene characteristics may constitute a separate unit of perceptual analysis and representation, which is at least as important and fundamental as that formed by the individual object.

In view of these considerations, several authors (Antes & Penland, 1981; Biederman, 1981; Friedman, 1979) have postulated scene-specific schemas or frames as the central notion in models of everyday object and scene perception. Figure 14.2 offers an outline of the resulting schema-driven view on real-world perception.

This theory holds that activation of a scene schema (i.e., a prototypical representation of the contents and spatial relations characteristic for a particular scene) mandatorily precedes identification of individual shapes and the completion of a full three-dimensional (3-D) parse of the scene. This is because such activation is based on low-resolution image information that can be processed extrafoveally and with unfocused attention (Metzger & Antes, 1983). In addition, it does not need to be analyzed up to a full 3-D level in order to allow access to a stored representation (Biederman et al., 1982). Both scene backgrounds (Boyce et al., 1989) and certain spatial configurations of prominent geons of scene-diagnostic objects (Biederman, 1988) have been argued to provide such information.

Following this rapid schema activation, it is assumed that individual object identification as well as the apprehension of interobject relations in the scene will be driven by the schema. It should be noted, however, that there is still debate as to how precisely the schema influences these processes. I indicated this in Figure 14.2 by outlining two different hypothetical influence routes (represented with broken lines).

The strongest position (e.g., Antes & Penland, 1981; Friedman, 1979) clearly is that schema activation makes it possible to directly achieve
identification of schema-consistent objects or object configurations by checking the image for the presence of gross 2-D and 2.5-D characteristics specified in the schema (influence route A in Figure 14.2). Overall shape, color, texture, dimensionality, aspect ratio, relative size, and position of image regions have all been suggested as relevant input for this concept-driven and resource-inexpensive mode of image interpretation. Schema-inconsistent or isolated objects, on the other hand, are identified on the basis of a data-driven and resource-expensive computation of a 3-D object model from the image, which is then compared to a detailed structural representation at the individual object level (left-hand side of Figure 14.2). Similarly, schema-inconsistent object configurations will have to be anchored in an image-driven 3-D scene parse before they can be tested against existing schemas in order to determine whether they should be marked as a deviation from an old scene or as an exemplar of an entirely novel scene.

More moderate than this postulation of two qualitatively different routes to recognition, is the view (e.g., Riseman & Hanson, 1987) that scene schemas are used as a basis for generating hypotheses about the pattern tests and image segmentations that are most likely to result in the construction of a 3-D model that is consistent both in terms of its internal structure and in terms of its relations to the rest of the scene. In this sense (influence route B in Figure 14.2), schema-influences do not alter the basis of visual recognition from data-driven pattern analysis to concept-driven pattern detection, but rather focus the image analysis system by constraining its search space.

Although this schema-driven alternative to the object-based view appears to be commonly accepted, I think its validity still needs to be questioned, both on theoretical and on empirical grounds. The central theoretical problem is clearly illustrated in the views on how scene-specific schemas are supposedly activated during the first glance at a scene. Antes, Mann, and Penland (1981) argued that such activation is the result of the rapid identification of one or more scene-diagnostic objects in the scene. They based this conclusion on the finding that the quality of global thematic scene characterizations significantly deteriorates when scene-diagnostic objects are either removed from the scene or replaced by "shapemates" (i.e., undetailed volumes with the same overall shape). Biederman (1981, 1988), however, argues against this route to schema activation and precisely claims that particular spatial configurations of these shapemates enable access to the appropriate schema. The main piece of evidence cited to support this claim is Mezzanotte's (1981) finding that intact target objects placed in a scene constructed entirely with shapemates, are more difficult to verify when they violate their normal spatial and semantic relations to that scene. Finally, Boyce et al. (1989) assign a crucial role to scene-backgrounds, on the basis of their observation that target verification is facilitated when the object is placed in a consistent background rather than in a nonsense background. On the one hand, this questions the exclusive importance of shapemate configurations in schema activation because the target was accompanied by the same set of objects (the target's cohort) in the same positions in both background conditions. On the other hand, it suggests that individual objects play no part at all in schema activation because neither target-cohort consistency nor cohort-background consistency had an effect on target verification.

This lack of a consensus on the modalities of rapid schema activation exemplifies the major theoretical flaw in present schema-driven views on
real-world perception, i.e., the *deus ex machina* fashion in which a scene-specific unit of perceptual analysis and representation is being postulated and the total absence of a coherent theory of scene recognition. Such a theory should answer questions about what defines a scene and distinguishes it from another scene, how tolerant recognition is with respect to deviations from these scene-diagnostic features and how this tolerance can be economically represented, how scene features can be extracted from images, what the representational format and contents of scene schemas need to be in order to be matched successfully with image information, how these representations have been learned, etc. It is clear, however, that answering these questions is a nontrivial enterprise which makes it very worthwhile to scrutinize the empirical necessity for doing away with the object-based view in favor of a schema-driven account.

The first, and best documented, finding that has been questioned as a valid basis for the schema-driven account, is the Probability Violation Effect (PVE), i.e., the decrease in ease of object identification when the object is presented in a scene it is unlikely to occur in (e.g., Antes & Penland, 1981; Loftus & Mackworth, 1978). The reason for this questioning, is that a number of studies established very similar effects on the ease with which isolated objects were identified (Carr, McCauley, Sperber, & Parmelee, 1982; Henderson, Pollatsek, & Rayner, 1987; Huttenlocher & Kubicek, 1983; Kroll & Potter, 1984). Specifically, it was found that the prior presentation of a semantically related object reduced object identification times, which was interpreted as evidence for an automatic spreading of activation between individual object representations. To the extent that semantically related objects also tend to appear in the same real-world scenes, one could therefore argue that interobject priming rather than schema mediation is at the basis of the PVE. In this view, the PVE does not necessitate adoption of the complex model outlined in Figure 14.2 because it can be completely accounted for by the mere addition of an activation link between stored object representations in Figure 14.1.

However, it could be argued that there appear to be constraints on the interobject priming process that may invalidate it as an explanation of the PVE. Based on the Henderson et al. (1987) studies, I examined three such constraints (De Graef, 1990).

The first constraint I examined was the limitation of priming to strictly semantically related objects. Although it is not an unreasonable assumption that these objects also tend to appear in the same scenes, one can undoubtedly present an impressive list of semantically unrelated objects that are also likely to appear in the same scenes (e.g., a toilet and a razor in a bathroom, a fireplace and a television in a living room, etc.). Given the robustness of the Probability Violation Effect across different studies and sets of stimuli, this could be taken to imply that priming cannot be the basic mechanism underlying the effect. However, using arrays of visually isolated objects I found evidence for priming between objects that were strictly selected on the basis of their episodic relatedness, i.e., their likelihood of appearing in the same scene. Specifically, foveating an episodically related object was found to produce a significant 46-ms reduction in the first fixation duration for the next object, relative to situations in which unrelated or no objects at all were fixated first. To the extent that first fixation duration reflects object identification time (e.g., Friedman, 1979; Henderson, Pollatsek, & Rayner, 1989; Loftus & Mackworth, 1978), this finding supports the hypothesis that interobject priming could indeed explain the Probability Violation Effect in full scenes.

The second constraint I examined was the limitation of priming effects to controlled processing of the primed object or target. In all studies cited above, the priming effect was measured contingent upon foveation of the target in order to name or memorize it. Advocates of the schema-driven view, however, have also reported the Probability Violation Effect for very briefly presented objects at extrafoveal and a priori uncertain locations in complex scenes (e.g., Biederman, 1981; Boyce et al., 1989). To determine whether priming could account for this phenomenon, I presented subjects with 150-ms masked exposures of arrays of isolated objects. Subjects were instructed to fixate in the center of the display, in preparation for a forced-choice object recognition task following each trial. In this manner, priming effects on automatic target processing could be examined by analyzing recognition performance for peripherally located targets as a function of the information in the display's center, i.e., a target-related object, a target-unrelated object, or no object at all. This analysis revealed similar performances in the latter two cases and a significantly worse, chance-level performance in the related prime condition. Although this finding was demonstrated to confirm the existence of an episodic priming mechanism, it also revealed that such a mechanism could not facilitate preattentive perceptual processing of objects. Consequently, if such processing is indeed reflected in the accuracy with which extrafoveal objects are verified in briefly presented scenes, then the Biederman (1981) and Boyce et al. (1989) studies show that priming is insufficient to fully explain the Probability Violation Effect in scene perception.

A similar conclusion was reached with respect to the third constraint, i.e., the limitation of priming effects to situations in which the prime itself has been subjected to controlled processing. While examining the two previous constraints, I also investigated whether the size of any episodic priming effects varied as a function of the number of nonattended target-related objects in the display. This analysis revealed no significant effects of target-display consistency on either first fixation durations (controlled target processing) or performance in the forced-choice task (automatic target processing). This finding confirms the Henderson, Pollatsek, and
Rayner (1987, 1988) conclusion that priming is strictly conditional upon the relation between the target and the object fixated just prior to it. It suggests that automatic prime processing cannot affect the perception of related objects, which leads to the prediction that facilitation of probable object identification in scenes should be preceded by a period of selective attention to another probable object. This prediction, however, runs counter to reports of a Probability Violation Effect for targets presented in foveal vision throughout a 100-ms scene exposure (Klatsky et al., 1981), as well as for targets placed in extrafoveal vision during 150 ms exposures of scenes from which foveal objects had been systematically removed (Boyce et al., 1989). Again, the priming explanation of the Probability Violation Effect in scenes is questioned, provided one can assume that the results of these studies reflect an impact on perceptual object processing.

In fact, the validity of this assumption can safely be regarded as the crucial issue in examining the empirical necessity for adopting a schema-driven view on real-world perception. Indeed, the strongest evidence for this approach stems exclusively from studies that use the previously described speeded target verification task. Specifically, I am referring to the reports by Biederman and colleagues that performance in this task decreases when the objects at the cued position inappropriately defy gravity, or appear in unlikely sizes or positions relative to the rest of the scene. This finding suggests that the spatial structure inherent in real-world scenes provides a contextual definition of a set of relational object features (i.e., support, position, and size) that, from the very first scene fixation on, is taken into account during object perception. Clearly, the object-based view would be inadequate to deal with this phenomenon, and could not be made to do so by simply adding unlabeled interobject activation links as was proposed in order to deal with the Probability Violation Effect. Rather, the schema-driven model, centered around the notion of an integrated representation of the typical contents and spatial layout of a particular object-context configuration, would be the most appropriate framework to explain these Spatial Violation Effects.

However, De Graef, Christiaens, and d’Ydewalle (1990) have cautioned against this conclusion based on the argument that response speed and accuracy in the speeded target verification task may not measure the perceptibility of the object at the cued position, but rather may reflect the subject’s degree of uncertainty in postperceptually deciding whether this object was indeed the prenamed target. I will not repeat the precise rationale of our argument, but its main thrust is that the brief and extrafoveal presentation of the cued object rarely provides detailed image information about its structural features, and that subjects therefore will often have to resort to educated guesses about its identity. For this purpose they can draw upon image cues for global thematic scene identity (e.g., texture or color regions, background information such as fork and arrow vertices formed by adjoining walls) and on information about gross spatial properties of the cued object (e.g., the proportion of the scene’s visual angle occupied by the object, its distance to the scene’s ground plane, its nearness and position relative to other scene components). Based on a comparison between this information and their a priori knowledge about the prenamed target object, subjects can generate postperceptual guesses that can be demonstrated (De Graef et al., 1990) to lead to the exact response patterns that previously have been interpreted as reflecting Probability and Spatial Violation Effects on object perceptibility.

Because this argument obviously challenges the entire empirical basis for the schema-driven model of real-world perception, De Graef et al. (1990) attempted to test the model’s validity with a less disputable paradigm for measuring the context sensitivity of object perception. Specifically, our approach involved the presentation of line drawings of realistic scenes in which selected target-objects appeared in either a normal or an inappropriate relation to the rest of the scene. The scenes also included a variable number of “non-objects,” a notion we gratefully borrowed from Kroll and Potter (1984) and by which we refer to figures that resemble objects in terms of size range and the presence of a clear and closed part-structure, but which are completely meaningless otherwise. An example of our stimuli is provided in Figure 14.3. Our subjects’
task was to freely explore the scenes in order to count the number of non-objects. During this scene exploration, we recorded fixation times for the target-objects that the subjects incidentally fixated in their search for non-objects.

This paradigm has two main advantages. First, it presents a task in which subjects only have to scan the scene and determine if any object-like entities they come across correspond to a known object. The absence of viewing constraints or mnemonic requirements in this “object decision task” is quite important because it means that subjects are not encouraged to deliberately capitalize upon context in order to either compensate for a lack of image information on structural object features or to facilitate memory trace formation (Schank, 1982). Consequently, any context effects observed in this task can be considered to be mandatory rather than elicited. Second, the registration of eye movements under these conditions provides an unobtrusive, on-line measure of object perceptibility. Specifically, for each object fixated in the course of scene exploration, first fixation duration can be determined as a measure of object identification time.

With this paradigm, we examined three main questions. First, could we replicate the Probability and Spatial Violation Effects? To test this, we compared fixation times for objects undergoing relational violations (Violation conditions) to those for the same objects appearing in a normal relation to their context (Base condition). Second, would these effects appear from the very first scene fixation on, as predicted by the schema-driven model of real-world perception? This was investigated by analyzing differences in fixation times for the normal and violated objects as a function of the ordinal position of their first fixation in the entire fixation sequence recorded for the scene they appeared in. Third, what is the precise role of contextual information in object perception? Specifically, does it directly influence the ease with which an object is apprehended in an image? Or does it serve as a framework for testing the plausibility of the output of a modular, strictly data-driven analysis of structural object features? Although the former view can be conceptualized in a number of ways (De Graef et al., 1990), all these accounts imply that a facility component will be present in the context effect. This is in contrast to the “plausibility-checking view,” which holds that context is merely used to endorse or reject the output of the object-encoding process and consequently predicts that the only effect of context will be to delay or inhibit conclusive object identification in the case of relational inconsistencies. To test which of these views is most appropriate, we included a condition in which objects appeared out of scene context (i.e., in an array of visually isolated objects) and compared object fixation times in this Isolation condition to those in the Base condition.

An overview of the results relevant to the first two questions can be found in Table 14.1, which presents the mean first fixation durations (FFD) for the Base and the Violation conditions, as a function of the moment in scene exploration at which the target object was fixated for the first time. “Early” and “Late” Fixation Moments were distinguished by determining ordinal fixation number for each first fixation on a target, and using the median of this distribution as the cutoff point. As can be seen in this table, an analysis of Base-Violation differences did reveal longer fixation times for objects violating Probability, Position, and Size, but only when these objects were fixated at later stages of scene exploration. As can be seen in a comparison of adjusted condition means (in parentheses), these effects remained unchanged when incidental differences in physical object characteristics with a possible effect on object perceptibility (i.e., object camouflage and absolute object size) were filtered out in an analysis of covariance.

With respect to our third research question, the pattern of results in Table 14.1 seems to indicate that the observed Violation effects reflect pure inhibition rather than a lack of facilitation. This is suggested by the fact that their late appearance is entirely due to an increase of fixation times for the violated objects, whereas fixation times for normal objects do not decrease. There is, however, one consideration that cautions against this conclusion. Specifically, studies examining the overall evolution in fixation durations during the exploration of pictorial stimuli all report considerable fixation duration increases over time. Only data reported by Loftus (1983) and the Base-condition data in the present experiment reveal no such tendency. Interestingly, both these exceptions were observed using pictures of realistic scenes whereas all other studies used stimuli such as the Thematic Apperception Test (TAT) cards (Antes, 1974) and abstract paintings (Locher & Nodine, 1987). Consequently, the reason for the discrepancy in findings may very well be that coherent real-world scenes provide stronger contextual constraints on individual scene-component identification. This suggests that the absence of a significant fixation duration increase in the Base condition may perhaps be the result of two counteracting processes, i.e., a general tendency for fixation times to increase during picture exploration and
a gradually developing contextual facilitation of object identification. By the same token, the fixation duration increase in the Probability, Position, and Support conditions may (at least partly) reflect a lack of facilitation normally produced by good context, rather than pure inhibition by bad context.

Although this hypothesis obviously calls for further systematic investigation, it was supported in an analysis of Base-Isolation differences as a function of Fixation Moment. Specifically, this analysis showed that object fixation times in the isolation condition significantly increased as a function of time spent in exploring the arrays, leading to first fixation durations that were significantly longer than those in the Base condition.

Clearly, our results are at odds with the immediate context effects observed in the speeded target verification paradigm. One possible explanation for this discrepancy is that we have failed to probe the perceptual processes reflected in the speeded target verification results.

This explanation was originally proposed by Biederman et al. (1982), who argued that object fixation times are unlikely to reflect object encoding because they are much longer than the exposure durations typically required to achieve high levels of object recognition accuracy in tachistoscopic studies. One problem with this argument is that estimating the precise time-difference between these two measures is nearly impossible as it is highly dependent upon the methodology involved in obtaining the measures. This certainly is the case for the fixation time measure as demonstrated by the fact that an older and coarser eye movement recording methodology used by Friedman (1979) yielded first fixation durations that were up to 400 ms longer than ours. In addition, the method-dependency problem also holds for the exposure duration measure. Biederman and Ju (1988), for instance, report a number of experiments that show that object recognition accuracy at a given exposure duration clearly varies as a function of task, the presence of a mask, speed-accuracy trade-offs, etc. The fatal flaw, however, in the fixation time versus exposure duration argument, is of course that it fails to acknowledge the fact that only a portion of the fixation time is actually used for visual encoding. That this may correspond to a time period that falls well within the range of the 50 to 150 ms exposure durations used in tachistoscopic studies, has been nicely demonstrated in foveal masking studies of reading processes (Rayner, Inhoff, Morrison, Sowizralak, & Bertera, 1981).

Another antifixation argument advanced by Biederman et al. (1982) is that violation effects on object fixation times are far too large to reflect increased encoding difficulties, but rather show a greater interest for “odd” objects. Again, basing the argument on absolute effect size is quite problematic because the Probability Violation Effect Biederman et al. refer to (Friedman, 1979) is about six times larger than the one we found. In addition, we did examine our data for a possible reflection of increased interest for violated objects. Specifically, due to the within-subjects nature of our experiment we were able to analyze object fixation times as a function of the number of times subjects had seen them violating normal object-context relations. Obviously, if longer fixation times for violated objects were due to greater interest, one would expect these fixation times to decrease as more violations were encountered and their interest value decreased. Our analyses showed that such an effect was indeed present for gaze durations (i.e., the total sum of successive fixations on an object), but not for first fixation durations. Furthermore, it should be noted that our controls for confounding influences of physical object characteristics showed a large effect of object camouflage on first fixation durations, which clearly is a desirable feature for a measure that is supposed to tap into visual processing. In this context, it is interesting to note that response speed and accuracy in the speeded target verification studies by Biederman and colleagues were completely unaffected by a manipulation (i.e., the “Interposition Violation”) in which the cued object’s featural structure was thoroughly disrupted by letting its background pass through it.

In view of these considerations, I feel rather confident about the validity of our paradigm and suggest that the observed discrepancy in results may be explained by reversing the argument. Do responses obtained with the speeded target verification paradigm reflect mandatory perceptual effects of context, or do they probe a deliberate, post-perceptual use of low-resolution image information as cues for making inferences about global scene and individual object identity? In this view, the speeded target verification results appear to be most relevant for developing algorithms to constrain the search space of man-made pattern recognition systems with a relatively limited capacity for data-driven image analysis. In models of human perception, however, these results do not necessitate the postulation of an immediately operating scene-encompassing mechanism for image interpretation, which mandatorily drives further object and scene identification. Instead, the perceptual impact of contextual information seems to develop only gradually, which suggests that it requires data-driven encoding of local scene components and the spatial relations between them.

Integrating the above arguments and findings, I think the framework outlined in Figure 14.4 may be most appropriate to capture the present status of research on scene-context effects and its implications for accounts of real-world perception. In essence, a data-driven view on real-world perception is proposed that preserves most of the main features of the object-based model. The outcome of object encoding and scene parsing is entered in a situational model of the viewed scene, which serves as a fluid interface between the image and general world knowledge. Parallel to its gradual definition, this model will start to influence further image processing. Our data suggest that this influence may be on
the actual encoding of the image, but much remains to be determined with respect to the exact locus of impact. Hence, the interrupted lines that represent hypothetical routes and mechanisms that may underlie the observed effects of Probability, Support, and Position violations. Although a separate scene-level of perceptual processing is considered unnecessary in the framework, the double lines on the right-hand side of Figure 14.3 were added to represent a non-perceptual explanation of the scene-context effects reported in the speeded target verification studies. Specifically, it is assumed that these results are produced by active utilization of low-resolution image cues, allowing for a thematic scene identification that will guide or completely replace the data-driven development of an ad hoc situation model.

Obviously, this proposal requires further development and will have to be amended in the face of new data. In fact, other chapters in the present volume (Boyce & Pollatsek; Henderson) already contain urgent calls for such development in their confirmation of our suspicion (De Graef et al., 1990) that differences in attentional distribution may mediate the presence of immediate and delayed components in the effects of context on perception. The presented view on scene-context effects is merely a step toward identifying the constraints that explain the ease with which we perceive our environment in spite of the tremendous computational complexities involved in even the simplest image interpretation tasks.

Acknowledgment. The writing of this chapter was supported by the Belgian government through agreement RFO/A1/04 of the Incentive Program for Fundamental Research in Artificial Intelligence.

References


