Dynamical approaches to cognitive science

Randall D. Beer

Dynamical ideas are beginning to have a major impact on cognitive science, from foundational debates to daily practice. In this article, I review three contrasting examples of work in this area that address the lexical and grammatical structure of language, Piaget's classic 'A-not-B' error, and active categorical perception in an embodied, situated agent. From these three examples, I then attempt to articulate the major differences between dynamical approaches and more traditional symbolic and connectionist approaches. Although the three models reviewed here vary considerably in their details, they share a focus on the unfolding trajectory of a system's state and the internal and external forces that shape this trajectory, rather than the representational content of its constituent states or the underlying physical mechanisms that instantiate the dynamics. In some work, this dynamical viewpoint is augmented with a situated and embodied perspective on cognition, forming a promising unified theoretical framework for cognitive science broadly construed.

The conceptual framework that we bring to the study of cognition can have profound empirical consequences on the practice of cognitive science. It influences the phenomena we choose to study, the questions we ask about these phenomena, the experiments we perform, and the ways in which we interpret the results of these experiments. Until relatively recently, there was 'only one game in town'[1] – the computational hypothesis that underlying cognition is the purely formal manipulation of quasi-linguistic symbolic representations by syntactic rules[1–3]. However, in the mid-1980s, the theoretical imagination of cognitive science was significantly expanded by the proliferation of connectionist models[4–6]. More recently, there has been a growing interest in dynamical approaches to cognitive science[7–12]. Drawing upon the mathematical tools of dynamical systems theory (see Box 1), a dynamical analysis of a cognitive process seeks to understand the unfolding of that process over time and the multiple internal and external influences whose interplay shapes this unfolding. In this article, I will briefly review three rather different examples of this dynamical approach to cognition. I will then compare and contrast this approach with the more traditional symbolic and connectionist approaches, and briefly discuss some of the contributions that dynamical ideas are making to ongoing debates about the foundations of cognitive science.

The lexical and grammatical structure of language

Language, that quintessentially human cognitive skill, is often seen as the strongest argument against non-symbolic approaches to cognition[13,14]. However, despite its symbolic character, language understanding and production are essentially temporal events, with preceding words strongly influencing the interpretation or selection of later ones. For this reason, language has been a major focus of dynamical modeling, ranging from models of the temporal structure of speech perception and production[15,16] to more abstract models of grammatical structure[17–19]. A good illustration of the development of a connectionist approach to language in which dynamical notions play a fundamental role is provided by the work of Elman[20]. Elman's models utilize simple recurrent networks[21], a variation of multilayer feedforward networks in which the hidden layer also receives inputs from a set of context units that hold the state of the hidden units from the previous time step. Such networks are non-autonomous discrete-time dynamical systems (see Box 1).

In one experiment, Elman used backpropagation to train a simple recurrent network to predict subsequent words in a sentence using a corpus of 10 000 short sentences formed from a vocabulary of 29 nouns and verbs[21]. Predicting word order requires the network to learn the frequency of occurrence of each possible successor word for each possible context, a task that demands a memory of previously seen words and an understanding of the grammatical structure of language. After training, a hierarchical cluster analysis of the hidden-unit activations for each word averaged across all contexts revealed that the trained network developed an internal dynamics whose organization reflected grammatical category and meaning. Specifically, it exhibited a major division into nouns and verbs, and subdivisions into animate and inanimate nouns and transitive, intransitive and optionally transitive verbs, and so on. However, a finer analysis that did not average over context revealed small but systematic differences between different senses and roles of words (e.g. 'boy' as subject versus 'boy' as object). Thus, a network state did not correspond to a word *per se*, as a traditional representational

R.D. Beer is at the Department of Electrical Engineering and Computer Science, and the Department of Biology, Case Western Reserve University, Cleveland, OH 44106, USA.

tel: +1 216 368 2816
fax: +1 216 368 2801
e-mail: beer@
eecs.cwru.edu
http://vorlon.cwru.
edu/~beer

Box 1. A brief introduction to dynamical systems theory

Because the terminology of dynamical systems theory is likely to be unfamiliar to many cognitive scientists, I will briefly review it here. (More comprehensive but still novice-friendly introductions can be found in Refs a and b.)

The concept of a dynamical system is very general. Loosely speaking, a dynamical system is a mathematical object that unambiguously describes how the state of some system evolves over time. More formally, a dynamical system is a triple $<T, S, \phi_t>$ consisting of an ordered time set T, a state space S, and an evolution operator $\phi_t: S \rightarrow S$ that transforms an initial state $x_0 \in S$ at time $t_0 \in T$ to

another state $x_t \in S$ at time $t \in T$. The time set T may be continuous or discrete. The state space S may be numerical or symbolic, continuous or discrete or a hybrid of the two, and it may be finite- or infinite-dimensional depending on the number of variables required to fully describe the state of the system. ϕ_t may be given explicitly or defined implicitly, it may be deterministic or stochastic and it may have inputs (non-autonomous) or not (autonomous). Sets of differential or difference equations, cellular automata, finite state machines and Turing machines are all examples of dynamical systems. Dynamical systems theory offers a
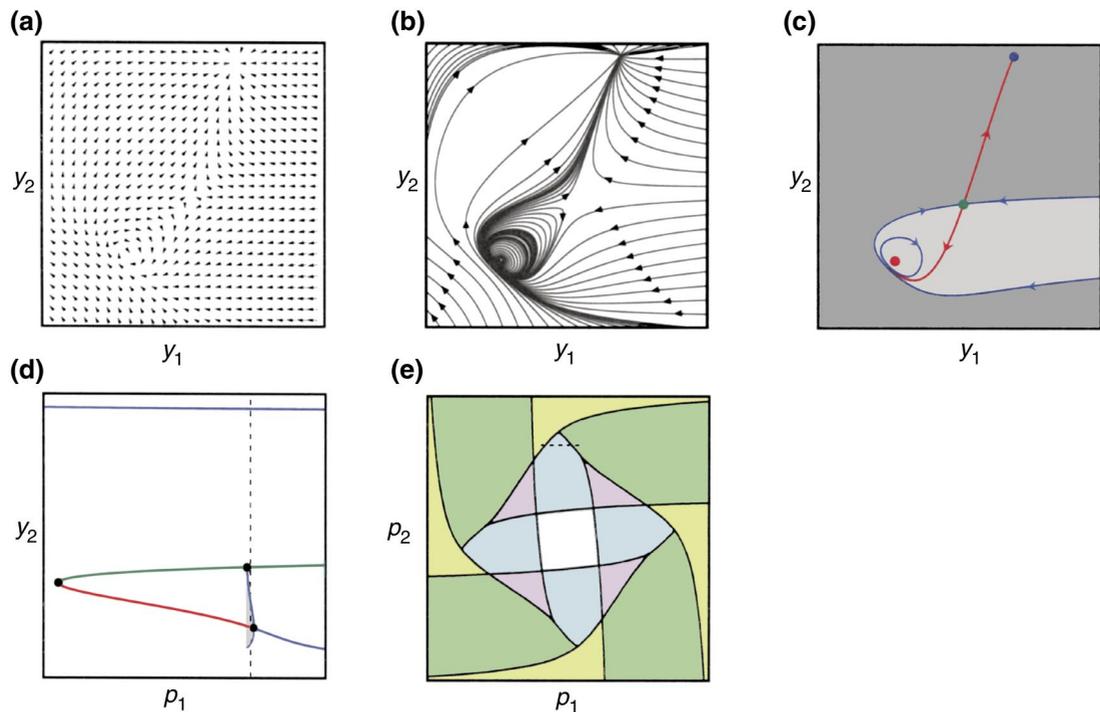
Fig. I. **Some basic concepts from dynamical systems theory.** **(a)** A vector field showing the instantaneous direction and magnitude of change at each point in the state space. **(b)** A flow showing representative solution trajectories of the system. **(c)** A phase portrait showing the limit sets, their stabilities, and their basins of attraction. Dots denote equilibrium points and the circular blue trajectory is a limit cycle. Stable limit sets are colored blue, unstable limit sets are red and saddle limit sets are green. The other blue trajectory and the red trajectory correspond to the stable and unstable manifolds, respectively, of the saddle point. The stable manifold of the saddle point separates the basin of attraction of the limit cycle (light gray) from the basin of attraction of the stable equilibrium point (dark gray). **(d)** A bifurcation diagram showing how the y_2 projection of the phase portrait changes as a parameter p_1 is varied. The color conventions are the same as in (c) except that bifurcation points are shown as black dots. The parameter slice corresponding to the phase portrait in (c) is shown as a dashed line. **(e)** A parameter chart showing regions of parameter space with distinct phase portraits as two parameters (p_1, p_2) are varied. Identically colored regions have qualitatively similar phase portraits. The parameter slice corresponding to the bifurcation diagram in (d) is shown as a dashed line.

---

analysis might expect, but rather to the outcome of processing a word within a particular context.

In a second experiment, Elman extended this work to longer and more complex sentences with long-distance dependencies involving number agreement, verb argument structure, and relative clauses[22]. The dynamics of a network trained on the prediction task was then examined by plotting projections of the trajectories of hidden-unit activation produced by sample sentences. The trajectory produced by the sentence '*Boy who chases boy chases boy*' is shown in Fig. 1. Note that occurrences of the same word at different points in the sentence leave the network in different states. These differences in network state correspond to the network's memory of the information required to process long-distance dependencies correctly. Because the local dynamics at each point determine the effect that subsequent words can have on

the network state, grammatical constraints are manifested in the structure of the network dynamics itself. Thus, as a sentence is processed, each word drives the network along one of the different trajectories allowed by the dynamics at that point, with context manifested as variations in state that influence subsequent processing.

The network can also process more deeply embedded clauses. Interestingly, its performance degrades with embedding depth more quickly on centered-embedded sentences (e.g. '*Witch that tiger that tinman hears sees tames lion*') than on right-branching sentences (e.g. '*Tinman hears tiger that sees witch that tames lion*'), a pattern that is also observed in human language understanding[23]. More recent work has explored the linguistic capabilities of recurrent networks in greater depth and has provided a more detailed understanding of how dynamics can be harnessed to solve language problems[24].

wide variety of tools for visualizing and analysing the temporal behavior of such systems.

Consider a particular two-dimensional continuous-time dynamical system from Ref. c, where ϕ_t is given implicitly by two autonomous differential equations $\dot{y}_1 = f_1(y_1, y_2)$ and $\dot{y}_2 = f_2(y_1, y_2)$. These two differential equations define a **vector field**, which assigns an instantaneous direction and magnitude of change to each point in the state space (Fig. Ia). Starting from some initial state, the sequence of states generated by the action of the dynamics is called a **solution trajectory**. Such a trajectory has the property that its tangent at each point is given by the vector field at that point. The set of all possible solution trajectories is called the **flow**, which graphically illustrates the action of the evolution operator ϕ_t (Fig. Ib). Unfortunately, explicit expressions for the solutions to most nonlinear sets of differential equations are not available. It was the great insight of Henri Poincaré, the father of the modern qualitative or geometric theory of dynamical systems, that a great deal of information about the flow of a dynamical system can be extracted without having an explicit expression for its solution trajectories.

Of particular interest is the possible long-term behavior of a dynamical system. Over time, the state of many dynamical systems eventually ends up in a small subset of the state space called a **limit set**. A limit set is invariant with respect to the dynamics, so that if the system's state ever falls on a limit set, the dynamics will act to keep it there indefinitely. Two simple types of limits sets are **equilibrium points** and **limit cycles**. An equilibrium point is a limit set consisting of a single point, producing a constant behavior (dots in Fig. Ic), while a limit cycle is a trajectory that closes on itself, producing an endless rhythmic behavior (closed blue curve in Fig. Ic). For **stable** limit sets or **attractors**, all nearby trajectories converge to the limit set, so that small perturbations away from the limit set will return there (blue equilibrium point and blue limit cycle in Fig. Ic). In contrast, any perturbation from an **unstable** limit set will not return to that limit set, but will instead be carried elsewhere by the dynamics (red equilibrium point in Fig. Ic). Finally, **saddle** limit sets are generally unstable, but have some special directions that are stable (green equilibrium point in Fig. Ic; the stable directions are indicated by the incoming blue curves).

In general, a dynamical system will possess multiple limit sets, with each attractor surrounded by a set of points that converge to it over time. This set of points is called its **basin of attraction**. In Fig. Ic, the dark gray region is the basin of attraction of the blue equilibrium point and the light gray region is the basin of attraction of the blue limit cycle. Note how the two trajectories near the middle of the right side of Fig. Ib end up in very different final locations. This is because the upper trajectory begins in the basin of attraction of the blue equilibrium point, while the lower trajectory begins in the basin of attraction of the limit cycle. A global characterization of all the limit sets of a dynamical system and their stabilities and basins of attraction is called a **phase portrait** (Fig. Ic). Such a portrait summarizes the different dynamical behaviors that a system can exhibit and where in its state space these different behaviors can be found.

If the flow depends on parameters, then it will change as those parameters are varied. In general, the flow will change smoothly. However, at **bifurcation** points, the topological type of the flow can change drastically even as the parameters are smoothly varied, with limit sets appearing, disappearing, or changing their stability. A **bifurcation diagram** illustrates how the phase portrait of a dynamical system depends on a parameter (Fig. Id). Here bifurcation points are shown as black dots. At small values of p_1, only a single stable equilibrium point occurs (upper blue line). As p_1 increases, a pair of equilibrium points appear, one unstable (red curve) and one saddle (green cruve). At still larger values of p_1, a stable limit cycle appears (gray region bounded by blue curve). This is the situation corresponding to Fig. Ic. Finally, as p_1 is increased even further, the size of the limit cycle shrinks to zero, leaving two stable equilibrium points separated by a saddle point. If two parameters are varied simultaneously, then regions of parameter space that exhibit different phase portraits may exist, and the boundaries between these distinct regions correspond to bifurcations. A map of the layout of these regions in parameter space and the boundaries between them is called a **parameter chart** (Fig. Ie). Together, bifurcation diagrams and parameter charts describe the range of dynamical behavior that a system is capable of producing and how that behavior changes as parameters are varied.

References

a Abraham, R.H. and Shaw, C.D. (1992) *Dynamics: The Geometry of Behavior*, Addison–Wesley

b Strogatz, S.H. (1994) *Nonlinear Dynamics and Chaos*, Addison–Wesley

c Beer, R.D (1995) On the dynamics of small continuous-time recurrent neural networks. *Adapt. Behav.* 3, 469–509

The A-not-B error in infant reaching

Work on motor behavior, especially within the ecological psychology tradition, has a long history of dynamical thinking[9,25–29], and some of this work has begun to probe developmental questions of strong cognitive interest[7]. For example, Thelen, Schöner, Scheier and Smith[30] have developed a dynamical model of Piaget's classic 'A-not-B' error in infants of 7–12 months of age[31]. In this task, an infant is faced with two similar opaque containers with lids (Fig. 2). Initially, an infant is trained to reach reliably for an object hidden in container A shortly after they observe it being hidden there. If the object is now hidden in container B in full view of the infant and a short delay is imposed, the majority of 7–12-month-old infants will still attempt to retrieve the object from container A. Piaget originally interpreted this error in terms of an immature concept of object permanence. However, numerous experimental studies have demonstrated that this error is remarkably sensitive to context[32]. For example, infants are less likely to make the A-not-B error if (1) the two containers are made visually distinct, (2) there is no delay between hiding and reaching, (3) the infant is highly interested in the object being hidden (e.g. a cookie), or (4) the infant's posture changes between the A trials and the B trial. Another intriguing observation is that fewer infants make the A-not-B error if they are only required to look to the correct container rather than reach for it. Such patterns of context sensitivity must be explained by any contending theory of the A-not-B error.

In order to account for these observations, Smith and Thelen[7] sketched a dynamical model that was subsequently fleshed out in collaboration with Schöner and Scheier[30] using a general dynamical theory of motor programming[33]. The core of the model is a one-dimensional field that gives the
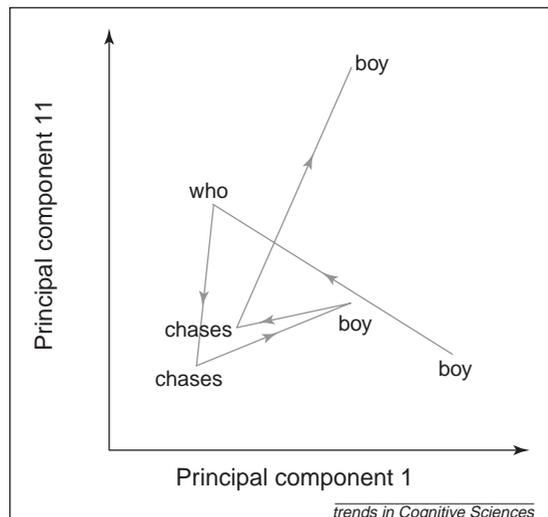
Fig. 1. A projection of the trajectory of a word prediction network as the sentence 'boy who chases boy chases boy' is processed. As each word is presented to the network, the state of activation of each of the hidden units changes. However, because of the large number of hidden units (70 in this case), this trajectory of hidden-unit activation cannot be directly plotted. Instead, a suitable two-dimensional projection is found using a technique known as 'principal components analysis' (PCA). By rotating the original coordinate axes, PCA finds a new set of axes which are ordered according to the amount of variance they account for. As shown here, projecting the 70-dimensional activation trajectory down to two principle components is sufficient to visualize the basic response. Note that occurrences of the same word in different contexts within the sentence leave the network in different states, which in turn affect the network's response to subsequent words. (Modified from Ref. 20.)

probability of reaching in a given direction. This 'movement planning field' (Fig. 2, light blue) receives two inputs describing the visual appearance of the scene: a constant 'task input' with peaks at A and B representing the two containers (Fig. 2, green), and a 'specific input' with a transient peak at either A or B representing the visual cue used to draw the infant's attention to the target object (Fig. 2, red). In addition, the movement planning field receives input from a one-dimensional 'memory field' that maintains a memory of recent reaches (Fig. 2, dark blue). The movement planning field also has a resting level that affects its ability to generate self-sustaining activity, as well as a source of noise. The model is thus an infinite-dimensional, continuous-time, stochastic, non-autonomous dynamical system (it is infinite-dimensional because its state is described by two continuous one-dimensional fields rather than a discrete set of state variables).

When the resting level of the field is low, the model produces the A-not-B error (Fig. 2, non-cooperative regime). In order to model the effects of training, the task input has an initial bias favoring A that decreases in subsequent trials. Initially, the model is exposed to repeated A trials by cueing it with a specific input with a peak at A. This produces a strong tendency to reach to A in the movement planning field and a memory of previous reaches to A in the memory field. Then, the model is cued with a specific input with a peak at B and a delay is imposed. Although the movement planning field initially develops a peak at B, this peak quickly decays over time, and the memory of earlier reaches to A dominates. Interestingly, when the resting level of the movement

planning field is high (Fig. 2, cooperative regime), the initial peak at B does not decay and the error is not produced.

Thus, this model suggests that the A-not-B error is due to an immature goal-directed reaching system rather than an immature concept of object permanence. On this dynamical account, the A-not-B error arises from the inability of the movement planning field to sustain a visually cued reach in a novel direction in the presence of a sufficiently strong memory of previous reaches. This suggests that the ability of the movement planning field to generate self-sustaining activity might be an important developmental parameter. The model can also account for the observed contextual effects (e.g. no error occurs when the delay between visual cue and reach is removed because the initial peak at B does not have to be sustained). It also makes several novel predictions (e.g. if the error arises from the general properties of goal-directed reaching, then it should be possible to observe it in older children as well if the visual cue is sufficiently weak and brief, the visual scene is sufficiently ambiguous, and the delay is sufficiently long).

Active categorical perception in an evolved model agent
Dynamical approaches in autonomous agents and robotics have also been an active area of research over the past ten years[34–38]. Of particular interest to the present discussion is work in which an evolutionary approach to the design of dynamical 'nervous systems' for model agents[34,36] is applied to the study of 'minimally cognitive behavior' (the simplest behavior that begins to raise questions of cognitive interest[39–41]). The dynamics of the evolved agents are then analyzed[10,42]. Studying simpler models that exhibit the basic features of a situated cognitive agent, but that are amenable to detailed experimental and theoretical analysis, is a powerful strategy for exploring the implications of a dynamical perspective[43].

In one study, a model agent was evolved to visually discriminate between two classes of objects of the same size, specifically, to catch circles and avoid diamonds[39]. The agent could move back and forth along a horizontal line while objects fell towards it from above (Fig. 3a). Using seven rays, the agent could sense the distance to each point of intersection between a ray and a falling object. The agent was controlled by a 14-neuron, continuous-time recurrent neural network[44]. The parameters of the neural network were set by an evolutionary algorithm[45] so as to minimize the final horizontal separation for circles and maximize the final horizontal separation for diamonds.

On this task, the best agent achieved an average performance of 99% on 100 random trials. The strategy that this agent used was first to foveate and then actively scan any object that appeared in its field of view. As the object neared, the agent would either decrease the amplitude of the scan until the object was centered (if it was a circle) or make a large avoidance movement (if it was a diamond) (results shown in Fig. 3b). Given the coarse resolution of the agent's seven rays, it is likely that active scanning accentuates the small differences between a circle and a diamond. Interestingly, active control of gaze direction has become an important theme in vision research[46,47]. Probing the agent with a class of objects that smoothly interpolated between circles and diamonds revealed a sigmoidal classification curve with a relatively sharp transition, a classic characteristic of categorical perception[48]. Experiments

trends in Cognitive Sciences

Fig. 2. B trial of the A-not-B task. The three main stages of a trial are illustrated at the top and corresponding plots of the state of the various fields in the model at each stage are shown below. The plots give the level of activity of each field as a function of horizontal position in the workspace: task input field (green), specific input field (red), memory field (dark blue) and movement planning field (light blue). Initially, the model is exposed to repeated A trials (not shown), resulting in a memory of previous reaches to A (as indicated by the peak at A in the left dark blue memory field plot). Then the B trial begins. First, the object (yellow star) is hidden in container B in full sight of the infant. Second, a delay is imposed. Third, the apparatus is pushed towards the infant and the infant reaches to one of the containers to re-trieve the object. When the resting level of the movement planning field is low (non-cooperative regime), a high initial peak at B decays over time (asterisks) to a reach to A and the model produces the A-not-B error. When the resting level of the movement planning field is high (cooperative regime), it is able to sustain the peak at B (asterisks) and the model produces a correct reach to B. (Modified from Ref. 30.)

with a variety of objects demonstrated that object width was the primary feature used in the discrimination.

How does the dynamics of the evolved neural network allow this agent to achieve such an accurate discrimination with such coarse sensors? The neural circuit is a 14-dimensional, non-autonomous, continuous-time dynamical system, while the entire coupled system comprising the neural circuit, the agent's body and the object is a 16-dimensional, autono-mous, hybrid dynamical system. Although it is impossible to visualize directly such high-dimensional dynamics, carefully chosen projections can provide important insights.

For example, we can understand how the agent's behavior arises from the interaction between the movement of its body, the movement of the object, and the dynamics of the evolved neural circuit by superimposing plots of the agent's motion on a depiction of the effects of the network dynamics on
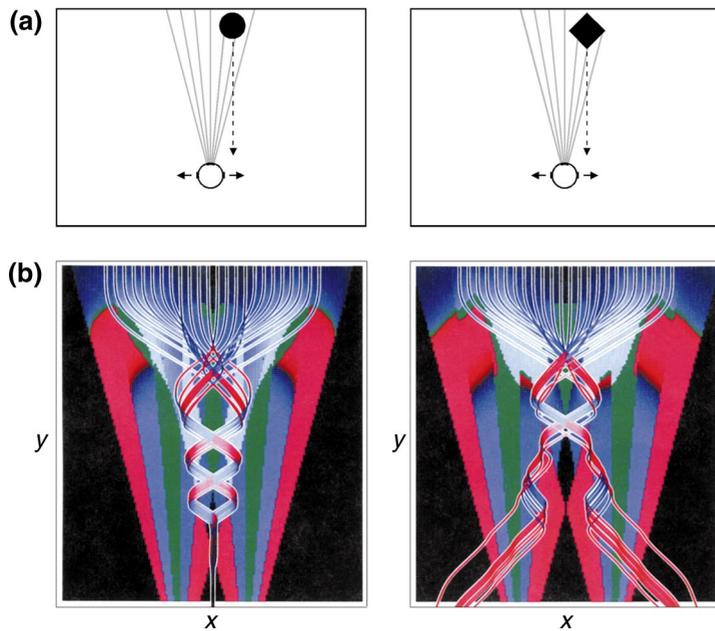
Fig. 3. Dynamical analysis of an evolved model agent for catching circles and avoiding diamonds. (a) The basic scenario. The agent can move back and forth horizontally while objects fall towards it from above. It uses an array of seven sensors to sense the distance to an object. **(b)** The movement of an evolved model agent for object discrimination on 24 circle trials (left) and 24 diamond trials (right). Here x is the horizontal position of the object relative to the agent (with zero at the center) and y is the relative vertical position (with zero at the bottom). The trajectories are color-coded according to the instantaneous horizontal velocity, with blue corresponding to a centering motion and red corresponding to an avoidance motion. The intensity of the color is proportional to the magnitude of the velocity, with black corresponding to a velocity of zero. These trajectories are superimposed on steady-state horizontal velocity fields, which show what the long-term horizontal velocity of the agent would be if an object were fixed at each point in its field of view. The steady-state velocity fields are color-coded in the same way as the trajectories, except that green indicates areas where the dynamics are multistable. In these cases, the steady-state velocity adopted depends on which basin of attraction (see Box 1) the transient trajectory finds itself within. Note that if an object were frozen at any point along a trajectory, the color of the trajectory would approach the color of the steady-state velocity field at that point over time.

ter until the object becomes trapped in the central black region of no movement and is caught. On the other hand, for diamonds the central red avoidance region is much larger and penetrates much higher in the agent's field of view, ultimately pushing the trajectories away from the center until the object is avoided.

Dynamical, symbolic and connectionist approaches to cognition

How does a dynamical approach to cognition differ from the more familiar symbolic and connectionist approaches? It might be argued that to distinguish dynamical approaches from these other approaches is mistaken for any of the following reasons: dynamical models can be simulated on digital computers[49], dynamical models can be given a computational description[50], Turing machines can be described as discrete-time dynamical systems over the integers[51], dynamical models can be represented as connectionist networks[52], and at least some connectionist networks are dynamical systems[6]. However, this is a bit like arguing that there is no difference between classical and relativistic mechanics because both are expressed using differential equations and both can be simulated on a digital computer. Clearly the important issues in that case were which differential equations were proposed, what theoretical entities those equations described, and, most importantly, the fundamentally different conceptions of space, time and the nature of the gravitational force that Newton and Einstein offered.

Likewise, regardless of the mathematical formalism or modeling technology employed, dynamical, symbolic and connectionist approaches differ markedly in the theoretical vocabulary and style of explanation that each brings to bear on cognitive phenomena[53]. A typical symbolic model is expressed as a program that takes as input a symbolic description of a problem to be solved. Then, using the system's general knowledge about the domain in which it operates (also symbolically represented), this description is manipulated in a purely syntactic fashion in order to obtain a solution to the problem[1–3]. Here, the explanatory focus is on the structure and content of the representations employed and the nature and efficiency of the algorithms used. Typical connectionist models, on the other hand, are expressed as layered networks of simple, neuron-like elements that are trained to transform a numerical input representation into a numerical output representation[4–6]. In this case, the explanatory focus is on the network architecture, the learning algorithm, and the intermediate distributed representations that are developed.

By contrast, a typical dynamical model is expressed as a set of differential or difference equations that describe how the system's state changes over time. Here, the explanatory focus is on the structure of the space of possible trajectories and the internal and external forces that shape the particular trajectory that unfolds over time, rather than on the physical nature of the underlying mechanisms that instantiate this dynamics. On this view, inputs do not uniquely specify an internal state that describes some external state of affairs. Rather, they serve as a source of perturbations to the system's intrinsic dynamics. For example, in the A-not-B model, the focus is on the response of the movement planning field to the perturbation provided by the transient specific input

that motion. Figure 3b shows the motion of a falling object through the agent's field of view as the agent moves back and forth while catching circles (left) or avoiding diamonds (right). Note how the initial states are separated into two distinct bundles of trajectories as the interaction proceeds. The trajectories are color-coded according to whether their instantaneous horizontal velocity is directed towards (blue) or away (red) from the centerline. These trajectories are superimposed on the steady-state horizontal velocity field, which shows the steady-state horizontal velocity generated by the neural circuit for an object fixed at each point in the agent's field of view. (These fields are color-coded in the same manner as the trajectories.) If the interaction were frozen at any point along a trajectory, then the instantaneous horizontal velocity would approach the steady-state value over time. However, as both the agent and the object are moving, the instantaneous velocity lags behind the steady-state value. Note how a red avoidance behavior often persists for some time after entering a blue, centering region, and vice versa. This subtle interplay between sensory input and internal state is crucial to accurate discrimination. For circles, the blue centering regions in the steady-state velocity field repeatedly turn the trajectory bundles back towards the cen-

(Fig. 2, red plots). Likewise, a system's internal state does not necessarily have any straightforward interpretation as a representation of an external state of affairs. Rather, at each instant in time, the internal state specifies the effects that a given perturbation can have on the unfolding trajectory. For example, in Elman's model, grammatical context is manifested by the differing effects that a word can have depending on the state of the network when the word is encountered (Fig. 1). The different states that result from encountering the same word in different contexts in turn differ in their response to subsequent words. In this way, an agent's past experiences influence its future interactions through its internal state, on multiple time scales. Thus, there are real conceptual differences between how a dynamicist, a computationalist and a connectionist approach a cognitive agent, differences that can have substantial empirical consequences in practice.

Although the major differences in outlook between dynamical, symbolic and connectionist approaches are clear enough, there is considerable fuzziness at the borders. For example, work with recurrent connectionist networks is clearly dynamical to the extent that it emphasizes the unfolding trajectory, rather than the nature and architecture of the underlying 'neuron-like' elements and their representational roles. More fundamentally, there is a great deal of controversy over exactly what a representation or a computation is. Some have argued that, although objectivist, symbolic notions of representation and computation might be inapplicable to dynamical models, suitably distributed, analog, context-dependent and action-oriented notions will be applicable[54–57]. However, great care must be taken in generalizing these notions. If any internal state is a representation and any systematic process is a computation, then a computational theory of mind loses its empirical force[58]. Regardless of how these debates are ultimately resolved, it is clear that dynamical ideas are forcing a critical evaluation of the notions of representation and computation within cognitive science.

Cognition and the dynamics of situated, embodied action

Although a dynamical approach can certainly stand alone, it is most powerful and distinctive when coupled with a situated, embodied perspective on cognition[59–63]. From this perspective, the principal aim of a situated agent is to take action appropriate to its circumstances and goals, and cognition is merely one resource among many in service of this objective. Other important resources include the physical properties of an agent's body, the structure of its immediate environment (including artifacts such as shopping lists, calendars and computers, etc.) and its social context. In this sense, cognition can extend beyond an agent's brain to be distributed over a system of people and objects within an environment[62,63]. A very strong parallel emphasis on situated and embodied action has also emerged in artificial intelligence and robotics[64–68].

In a dynamical approach to situated action, an agent's nervous system, its body and its environment are viewed as coupled dynamical systems[10] (Fig. 4). Given that bodies and nervous systems co-evolve with their environments, and only the behavior of complete animals is subjected to selection, the need for such a tightly coupled perspective should not be surprising. The focus here is on continuously engaging an environment with a body so as to stabilize appropriate

trends in Cognitive Sciences

Fig. 4. A dynamical perspective on a situated, embodied agent. The nervous system, body and environment are each conceptualized as dynamical systems, which are in constant interaction with each other.
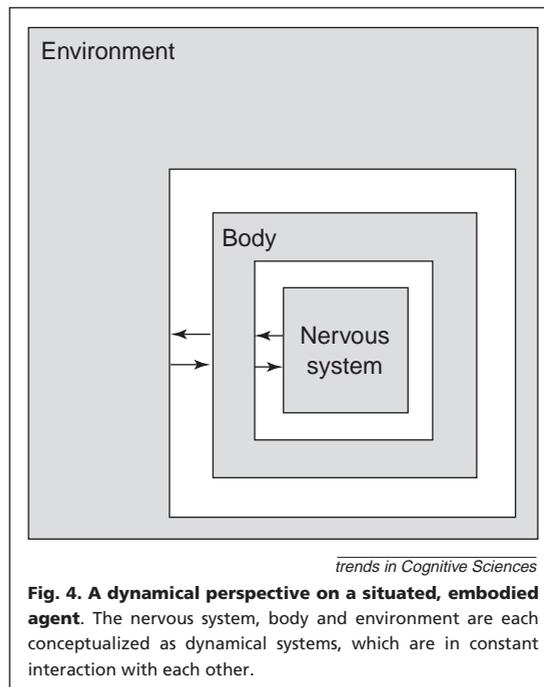
coordinated patterns of behavior, rather than the sequential sense–think–act processing cycle that is typical of computational approaches. For example, there is a very real sense in which the evolved model 'nervous system' for object discrimination described above does not itself know the difference between circles and diamonds. It is only when embodied in its particular body and situated in the environment in which it evolved that the distinction between circles and diamonds arises over time through the continuous interaction of these components. Because, as noted above, the internal state of the agent's dynamics plays such an essential role in structuring its behavior, this is no mere return to stimulus–response behaviorism. Rather, the working hypothesis of the dynamical approach is that, through increasingly sophisticated uses of internal state to mediate between perception and action, more cognitive behavior emerges from the dynamics of situated action. Thus, by supplying a common language for cognition, for the neurophysiological processes that support it, for non-cognitive human behavior, and for the adaptive behavior of simpler animals, a dynamical approach holds the promise of providing a unified theoretical framework for cognitive science, as well as an understanding of the emergence of cognition in development and evolution.

Conclusions

My primary goals in this review have been threefold. First, I have tried to communicate a sense of the practice of dynamical approaches to cognition by describing in some detail three rather different examples. These examples also demonstrate that dynamical approaches are beginning to engage substantive empirical questions in cognitive science. Second, I have argued that, although there may be a great deal of overlap between the modeling technologies and mathematical formalisms employed by dynamical, symbolic and connectionist approaches to cognition, there are very real differences in the conceptual frameworks they offer. Furthermore, these theoretical differences can have profound empirical consequences,

Outstanding questions

• Do dynamical, connectionist or symbolic conceptions of cognition offer the most powerful and parsimonious foundation for cognitive science, or will a hybrid of the three approaches be necessary?
• Are the existing mathematical tools and theoretical vocabulary of dynamical systems theory sufficient for applications in cognitive science, or will significant extensions be required?
• What is the proper formulation of representation and computation, if any, for situated, embodied dynamical agents?
• How will a dynamical approach to situated action fare on more paradigmatically cognitive phenomena, such as off-line deliberative reasoning and language understanding?

influencing the phenomena that are considered to be cognitive, the questions asked about these phenomena, the experiments performed, and the ways in which the results of these experiments are interpreted. Third, I have tried to show how, regardless of the eventual fate of the dynamical approach as an alternative or an adjunct to more traditional approaches, dynamical ideas are forcing a much-needed critical evaluation of the notions of representation and computation in cognitive science. They are also contributing to a more general broadening of cognitive science from its historically narrow focus on disembodied, language-like reasoning to embodied, situated action.

References

1 Fodor, J.A. (1975) *The Language of Thought*, Harvard University Press
2 Newell, A. and Simon, H.A. (1976) Computer science as empirical inquiry: symbols and search. *Commun. ACM* 19, 113–126
3 Pylyshyn, Z.W. (1984) *Computation and Cognition*, MIT Press
4 Rumelhart, D.E., McClelland, J.L. and The PDP Research Group (1986) *Parallel Distributed Processing, Vol. 1: Foundations*, MIT Press
5 McClelland, J.L., Rumelhart, D.E. and The PDP Research Group (1986) *Parallel Distributed Processing, Vol. 2: Psychological and Biological Models*, MIT Press
6 Smolensky, P. (1988) On the proper treatment of connectionism. *Behav. Brain Sci.* 11, 1–74
7 Thelen, E. and Smith, L.B. (1994) *A Dynamic Systems Approach to the Development of Cognition and Action*, MIT Press
8 van Gelder, T. (1995) What might cognition be, if not computation? *J. Philos.* XCI, 345–381
9 Kelso, J.A.S. (1995) *Dynamic Patterns*, MIT Press
10 Beer, R.D. (1995) A dynamical systems perspective on agent–environment interaction. *Artif. Intell.* 72, 173–215
11 Port, R.F. and van Gelder, T., eds (1995) *Mind as Motion*, MIT Press
12 van Gelder, T. (1998) The dynamical hypothesis in cognitive science. *Behav. Brain Sci.* 21, 615–665
13 Fodor, J.A. and Pylyshyn, Z.W. (1988) Connectionism and cognitive architecture: a critical analysis. *Cognition* 28, 3–71
14 Pinker, S. and Prince, A. (1988) On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73–193
15 Port, R.F. *et al.* (1995) Naive time, temporal patterns and human audition. In *Mind as Motion* (Port, R.F. and van Gelder, T., eds), pp. 339–371, MIT Press
16 Browman, C.P. and Goldstein, L. (1995) Dynamics and articulatory phonology. In *Mind as Motion* (Port, R.F. and van Gelder, T., eds), pp. 175–193, MIT Press
17 Pollack, J.B. (1991) The induction of dynamical recognizers. *Machine Learn.* 7, 227–252
18 Seigelmann, H.T. (1999) *Neural Networks and Analog Computation: Beyond the Turing Limit*, Springer-Verlag
19 Petitot, J. (1995) Morphodynamics and attractor syntax: constituency in visual perception and cognitive grammar. In *Mind as Motion* (Port, R.F. and van Gelder, T., eds), pp. 227–281, MIT Press
20 Elman, J.L. (1995) Language as a dynamical system. In *Mind as Motion* (Port, R.F. and van Gelder, T., eds), pp. 195–225, MIT Press
21 Elman, J.L. (1990) Finding structure in time. *Cognit. Sci.* 14, 179–211
22 Elman, J.L. (1991) Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learn.* 7, 195–225
23 Weckerly, J. and Elman, J.L. (1992) A PDP approach to processing center-embedded sentences. In *Proc. 14th Annu. Conf. Cognit. Sci. Soc.*, pp. 414–419, Erlbaum
24 Rodriguez, P. *et al.* (1999) A recurrent neural network that learns to count. *Connection Sci.* 11, 5–40
25 Kugler, P.N. and Turvey, M.T. (1987) *Information, Natural Law, and the Self-Assembly of Rhythmic Movement*, Erlbaum
26 Schöner, G. and Kelso, J.A.S. (1988) Dynamic pattern generation in behavioral and neural systems. *Science* 239, 1513–1520
27 Grossberg, S. and Kuperstein, M. (1989) *Neural Dynamics of Adaptive Sensory–Motor Control*, Pergamon Press
28 Bizzi, E. *et al.* (1991) Computations underlying the execution of movement: a biological perspective. *Science* 253, 287–291
29 Warren, W.H.J. (1995) Self motion: visual perception and visual control. In *Perception of Space and Motion* (Epstein, W. and Rogers, S., eds), pp. 263–325, Academic Press
30 Thelen, E. *et al.* The dynamics of embodiment: a field theory of infant perseverative reaching. *Behav. Brain Sci.* (in press)
31 Piaget, J. (1954) *The Construction of Reality in the Child*, MIT Press
32 Smith, L.B. *et al.* (1999) Knowing in the context of acting: the task dynamics of the A-not-B error. *Psychol. Rev.* 106, 235–260
33 Schöner, G. *et al.* (1997) The dynamic neural field theory of motor programming: arm and eye movements. In *Self-Organization, Computational Maps and Motor Control* (Morasso, P.G. and Sanguineti, V., eds), pp. 271–310, Elsevier
34 Beer, R.D. and Gallagher, J.C. (1992) Evolving dynamical neural networks for adaptive behavior. *Adapt. Behav.* 1, 91–122
35 Smithers, T. (1992) Taking eliminative materialism seriously: a methodology for autonomous systems research. In *Toward a Practice of Autonomous Systems* (Varela, F.J. and Bourgine, P., eds), pp. 31–40, MIT Press
36 Cliff, D. *et al.* (1993) Explorations in evolutionary robotics. *Adapt. Behav.* 2, 73–110
37 Schöner, G. *et al.* (1995) Dynamics of behavior: theory and applications for autonomous robot architectures. *Robotics Auton. Syst.* 16, 213–245
38 Tani, J. and Nolfi, S. (1999) Learning to perceive the world as articulated: an approach for hierarchical learning in sensory–motor systems. *Neural Netw.* 12, 1131–1141
39 Beer, R.D. (1996) Toward the evolution of dynamical neural networks for minimally cognitive behavior. In *From Animals to Animats 4: Proc. 4th Int. Conf. on Simulation of Adaptive Behavior* (Maes, P. *et al.*, eds), pp. 421–429, MIT Press
40 Parisi, D. (1997) Artificial life and higher level cognition. *Brain Cognit.* 34, 160–184
41 Gallagher, J.C. and Beer, R.D. (1999) Evolution and analysis of dynamical neural networks for agents integrating vision, locomotion and short-term memory. In *Proc. Gen. Evol. Comput. Conf.* (Banzhaf, W. *et al.*, eds), pp. 1273–1280, Morgan Kaufmann
42 Husbands, P. *et al.* (1995) Circle in the round: state space attractors for evolved sighted robots. *Robotics Auton. Syst.* 15, 83–106
43 Beer, R.D. (1997) The dynamics of adaptive behavior: a research program. *Robotics Auton. Syst.* 20, 257–289
44 Beer, R.D. (1995) On the dynamics of small continuous-time recurrent neural networks. *Adapt. Behav.* 3, 469–509
45 Mitchell, M. (1996) *An Introduction to Genetic Algorithms*, MIT Press
46 Ballard, D. (1991) Animate vision. *Artif. Intell.* 48, 57–86
47 Churchland, P.S. *et al.* (1994) A critique of pure vision. In *Large-Scale Neuronal Theories of the Brain* (Koch, C. and Davis, J., eds), pp. 23–60, MIT Press
48 Studdert-Kennedy, M. *et al.* (1970) Motor theory of speech perception:

a reply to Lane's critical review. *Psychol. Rev.* 77, 234–249

49 Garson, J.W. (1998) Why dynamical implementation matters. *Behav. Brain Sci.* 21, 641–642

50 Crutchfield, J.P. (1998) Dynamical embodiments of computation in cognitive processes. *Behav. Brain Sci.* 21, 635

51 Branicky, M.S. (1995) Universal computation and other capabilities of hybrid and continuous dynamical systems. *Theor. Comput. Sci.* 138, 67–100

52 Bridgeman, B. (1998) The dynamical model is a Perceptron. *Behav. Brain Sci.* 21, 631–632

53 Beer, R.D. (1998) Framing the debate between computational and dynamical approaches to cognitive science. *Behav. Brain Sci.* 21, 630

54 Clark, A. (1997) The dynamical challenge. *Cognit. Sci.* 21, 461–481

55 Mitchell, M. (1998) Theories of structure versus theories of change. *Behav. Brain Sci.* 21, 645–646

56 Wheeler, M. and Clark, A. (1999) Genic representation: reconciling content and causal complexity. *Br. J. Philos. Sci.* 50, 103–135

57 Clark, A. and Grush, R. (1999) Towards a cognitive robotics. *Adapt. Behav.* 7, 5–16

58 Beer, R.D. (1995) Computational and dynamical languages for autonomous agents. In *Mind as Motion* (Port, R.F. and van Gelder, T., eds), pp. 121–147, MIT Press

59 Winograd, T. and Flores, F. (1986) *Understanding Computers and Cognition*, Ablex

60 Suchman, L. (1987) *Plans and Situated Actions*, Cambridge University Press

61 Varela, F.J. *et al.* (1991) *The Embodied Mind*, MIT Press

62 Hutchins, E. (1995) *Cognition in the Wild*, MIT Press

63 Clark, A. (1997) *Being There*, MIT Press

64 Brooks, R.A. (1986) A robust layered control system for a mobile robot. *IEEE J. Robotics Autom.* RA-2, 14–23

65 Agre, P. and Chapman, D. (1987) Pengi: an implementation of a theory of activity. In *Proc. AAAI-87*, pp. 268–272

66 Kaelbling, L.P. and Rosenschein, S.J. (1990) Action and planning in embedded agents. *Robotics Auton. Syst.* 6, 35–48

67 Beer, R.D. (1990) *Intelligence as Adaptive Behavior*, Academic Press

68 Clancey, W.J. (1997) *Situated Cognition*, Cambridge University Press

Visually guided collision avoidance and collision achievement

David Regan and Robert Gray

To survive on today's highways, a driver must have highly developed skills in visually guided collision avoidance. To play such games as cricket, tennis or baseball demands accurate, precise and reliable collision achievement. This review discusses evidence that some of these tasks are performed by predicting where an object will be at some sharply defined instant, several hundred milliseconds in the future, while other tasks are performed by utilizing the fact that some of our motor actions change what we see in ways that obey lawful relationships, and can therefore be learned. Several monocular and binocular visual correlates of the direction of an object's motion relative to the observer's head have been derived theoretically, along with visual correlates of the time to collision with an approaching object. Although laboratory psychophysics can identify putative neural mechanisms by showing which of the known correlates are processed by the human visual system independently of other visual information, it is only field research on, for example, driving, aviation and sport that can show which visual cues are actually used in these activities. This article reviews this research and describes a general psychophysically based rational approach to the design of such field studies.

D. Regan is at the Departments of Psychology and Biology, York University, BSB, Room 375, 4700 Keele Street, North York, Ontario, Canada M3J 1P3. R. Gray is at Nissan Cambridge Basic Research, 4 Cambridge Center, Cambridge, MA 02142, USA.

**tel: +1 416 736 5627
fax: +1 416 736 5814
e-mail: dregan@
yorku.ca.**

W̲hat kinds of visual information are used by humans to avoid or achieve a collision? This question has driven a considerable body of research on a wide variety of topics, including highway safety, the design of flight simulators for training commercial airline pilots, and ballgames.

One approach to this issue breaks the problem into two parts: (1) how do we estimate the direction of an object's motion in depth; and (2) how do we estimate time to collision (TTC)? Boxes 1 and 2 describe relevant visual corre-

lates of these two quantities. Regarding the second part of the problem, many field studies seem to have been designed around the assumption that TTC is estimated entirely on the basis of 'tau' (Box 2). Wann[1] made the strong criticism that many such field studies did not effectively test the tau hypothesis or indeed any rival hypothesis; in effect, the tau hypothesis was regarded as an axiom rather than an hypothesis. Tresilian has recently provided a critical review of this sub-topic[2].