Essential title page information

Title. Physics-based character animation and human motor control

Author names and affiliations.
Joan Llobera *
Artanim Foundation 40, chemin du Grand-Puits
1217 Meyrin – Geneva, Switzerland
joan.llobera@artanim.ch

Caecilia Charbonnier
Artanim Foundation 40, chemin du Grand-Puits
1217 Meyrin – Geneva, Switzerland

* corresponding author

Highlights

- There are both theoretical and practical benefits to be found in a closer collaboration between research in Physics-based Character Animation (PBCA) and Human Motor Neuroscience.
- Modern physics-based animation controllers use deep reinforcement learning to learn how to apply torques to imitate reference animations.
- Some authors address interaction with kinematic controllers. In this case, the tendency is to train with large datasets.
- Other authors use specialized rewards to integrate interaction within the deep reinforcement learning policy tasked with the imitation of one or few reference animations.
- Highly constrained tasks can be learnt with no reference animations.
- Latent spaces allow skill transfer between tasks and compact representations of diverse motor actions.
- We review human motor control, including the roles of the cerebellum, the somatosensory cortex and the spinal cord.
- The role of the motor cortex in motor control has recently evolved away from the classical paradigm based on Penfield homunculus towards ethological activation maps.
- We propose to interpret the combined role of the motor cortex and the spinal cord as the encoder and decoder parts of a latent space of motor actions.
- More sophisticated forward models in PBCA controllers, similar to how the cerebellum enforces motor coordination in human motor control, may allow creating humanoid characters that collaborate significantly better with humans in a shared virtual reality or even in human-robot interaction scenarios.
- PBCA controllers can help validate empirically whether joint action paradigms as characterized in motor neuroscience work in practice. This in turn may bring practical benefits in the fields of human-character interaction and human-robot interaction.

Abstract :Motor neuroscience and physics-based character animation (PBCA) approach human and humanoid control from different perspectives. The primary goal of PBCA is to control the movement of a ragdoll (humanoid or animal) applying forces and torques within a physical simulation. The primary goal of motor neuroscience is to understand the contribution of different parts of the nervous system to generate coordinated movements. We review the functional principles and the functional anatomy of human motor control and the main strategies used in PBCA. We then explore common research points by discussing the functional anatomy and ongoing debates in motor neuroscience from the perspective of PBCA. We also suggest there are several benefits to be found in studying sensorimotor integration and human-character coordination through closer collaboration between these two fields.

Keywords: character animation; physics-based animation controllers; deep reinforcement learning; motor neuroscience; sensorimotor integration

Content

Physics-based character animation and human motor control

1. Introduction

Human and primate neuroscience have studied motor control for more than 100 years (Rosenbaum 2009). The overall objective of the field is understanding the biological mechanisms that support movement in humans and primates, using different methods from biomechanics, behavioural psychology, and motor neuroscience.

Physics-based character animation (PBCA) is a rapidly evolving sub-field of computer graphics, with a much shorter history. The primary goal of physics-based character animation is controlling the movement of an interactive character simulated as a ragdoll by applying forces and torques within a physics simulation. The control of the interactive character can be a goal by itself, or a preliminary step before applying these same forces in a robot. The targeted character is generally a humanoid, sometimes a quadruped, but other animals can also be mimicked, like a dolphin (Grzeszczuk, Terzopoulos, and Hinton 1998) or an ostrich (La Barbera et al. 2021). Two of the main open challenges in the field are, first, understanding the extent to which the movement synthesized can adjust interactively to real time perceptions or external input and, second, finding out how to control movement that involves rich contacts with the environment, like object manipulation.

These two research fields have developed quite independently. They often follow different methodologies and theoretical assumptions, and there are little cross-citations. However, it is easy to see convergent research objectives. Robotics and deep learning research are fields close to PBCA and have a rich tradition of exchanging ideas with human and primate motor control (Schaal and Schweighofer 2005; Floreano, Ijspeert, and Schaal 2014; Merel, Botvinick, and Wayne 2019). From this perspective, it seems reasonable that the functional characterization of how a human or a primate generates interactive behaviour can be modelled and validated in a simulation based on physics-based character animation techniques.

The main motivation for this review is to better understand the different perspectives available to study human-character interaction in virtual reality (VR) experiences. VR devices are now available as consumer hardware and allow rich and close interaction between humans and virtual characters. VR characters can engage with VR users, and we can study how well they manage to achieve cooperative or competitive tasks, compared to how 2 VR users do it (Llobera et al. 2022). Comparing characters controlled with PBCA techniques and humans doing individual or joint tasks in VR can cement new ground for the study of interpersonal coordination and, in general, of behavioural studies which require inter-person interaction. It can also bring new performance benchmarks for physics-based animation controllers. In this review we will consider the following aspects of human (and primate) motor control from the perspective of human neuroscience and PBCA:

1. Motor control integrates sensory input (to throw a ball to a target we need to know where the target is), and this sensory input produces significant behavioural variability.
2. Skill acquisition is goal driven and can work from few examples (compared to a large use of the skill acquired and a large variability of behaviour produced)
3. Emotion has a significant impact on behaviour, even in tasks non-related with the emotional state of the agent.
4. Joint action synchronisation is a widespread and ubiquitous phenomenon. It occurs spontaneously and it has significant social consequences (like bonding) and cognitive consequences (like improving learning).

Characteristics 1 and 2 are often considered in PBCA research, but characteristics 3 and 4 are generally disregarded. However, since these last two are of paramount importance in any task that involves coordination or social interaction, and our interest is in human-character interaction in VR, they are difficult to overlook. Given the broadness of the topic, we limit ourselves to humans (and primates, as a proxy to humans) and to physics-based humanoid interactive characters. We also limit our scope to short-term motor control, and do not discuss motor planning or language-related movements.

The document is organised as follows. In Section 2 we review the neuroscience of motor control. Section 2.1 reviews the extent to which principle-driven theories of motor function account for the four previous aspects in a coherent account of human (or primate) behaviour. In section 2.2 we explore in further detail the role associated with different parts of the nervous system involved in motor control, and to what extent this nuances the general theories introduced in section 2.1. We also use this section to introduce notions like latent spaces and forward models, which are ubiquitous in physics-based character animation.

In Section 3 we turn to PBCA. In section 3.1 we review the building blocks available: physics engines and actuators generally used. Section 3.2 introduces forward models in the context of character control. Section 3.3 reviews inverse models, which form most of recent contributions to physics-based character animation. We first introduce essential definitions involved in deep reinforcement learning, the method on which most of inverse control models are based. A taxonomy of these methods is then introduced, based on two different strategies to handle the interactive aspect of motor control. We then discuss, for each strategy, the implications for training datasets associated with adopting one or another strategy, what does “learning motor control” mean, and the difference in the training setup.

Section 4 reviews the field from a slightly more theoretical angle: can we characterize the space of motor actions? What is a distance between animations? We first review how this has been developed in PBCA, and then revisit the characterization of human motor control done in Section 2 from the perspective of PBCA research. We also comment the extent to which this perspective fits with ongoing debates on the organization of motor actions in the neuroscience literature. Section 5 discusses the open challenges involved in developing PBCA controllers that work better when engaging in cooperative tasks with humans, reviewing the four aspects of human motor control previously outlined and proposing possible directions forward. An appendix containing a glossary of PBCA terms is also added as an appendix to help the reader find more quickly the relevant terms.

2. The neuroscience of human motor control

2.1. Functional principles of human motor control

There are different principle-based theories to model and interpret human motor control. These theories use functional principles: their development is based on principles that can be formalized mathematically, and evidence for their validation is sought with histological, physiological and behavioural approaches. None of the approaches address completely the four aspects of motor control in which we are interested (see Table 1), but they form a good starting point to try explaining motor behaviour, and in this section we will review summarily the main existing ones. We also introduce the idea of a forward model and a latent space, notions that are useful to understand motor control, both in biological systems and in artificial humanoids.

2.1.1. Optimal Control Theory and Sensorimotor Integration

Sensorimotor integration studies how we adapt our movement according to particular sensory input. Examples can be how we reach differently for a glass, depending on where it is on a table, or how we adjust the swing of a bat to match an incoming ball. Principled approaches to sensorimotor integration have traditionally focused on insights derived from Optimal Control Theory (Bian, Wolpert, and Jiang 2020; Mathis and Schneider 2021). Optimal Control Theory allows reconciling how for a given task behavioural goals are completed reliably while there is often large variability in the detail of the movements performed to achieve it (Todorov and Jordan 2002; Todorov 2004; Scott 2004). The essential assumption is that goal completion under noisy perception conditions is what explains trajectory variability. Based on this assumption experimental predictions could be made, which were in turn validated experimentally with animals or humans.

The main success of optimal control theory is introducing a principled way to explain how challenging actions that involve sensorimotor integration can be achieved through training. An important and widespread assumption of Optimal Control Theory is that the brain has an internal model of sensorimotor action (Wolpert, Ghahramani, and Jordan 1995) and, in particular an *efferent copy* which is used to predict the outcome of an action before it is executed. Comparing the predicted outcome with the perceived outcome, and allows correcting the action to achieve the targeted goal (Miall and Wolpert 1996). This idea, the use of a **forward model** that, by predicting the outcome of performing a given behaviour, can help better control movement, is one that has found widespread developments across neuroscience, robotics, and character animation. More recent work combining Optimal Control with deep neural networks has also help map neural coding properties in the motor cortex (Lillicrap and Scott 2013).

Optimal Control Theory is sometimes presented as an alternative to PBCA to modelling sensorimotor control (Hausmann et al. 2021). However, the optimality assumption is also a limitation: it's still unclear the extent to which this modelling approach can apply to complex movements, or to movements beyond relatively short timeframes (Merel, Botvinick, and Wayne 2019). Even less for behaviours that do not have a clear optimality criterion. It also fails to explain why or how emotion affects our behaviour (Rosenbaum 2009) or it impacts the outcome of physical activities (Beedie, Terry, and Lane 2000). Therefore, research in sensorimotor integration can skip entirely the assumption of optimality and focus on analyzing how different parts involved in motor control interact (Asan, McIntosh, and Carmel 2022)

2.1.2.Free-energy minimization and Active Inference

Free-energy minimization has been used as an underlying principle to explain human skill acquisition in both perception (K. Friston 2009; Costa et al. 2020) and action (K. Friston 2010) including motor synthesis and motor planning (K. Friston et al. 2016). From this perspective, there is a central, global task in the brain, which is to minimise sensory prediction error. The minimization is achieved through Hebbian learning (i.e., a biological neural network, opposed to back propagation used in artificial neural networks). Active Inference is an extension of the idea of free-energy minimization which proposed that the motor system acts in a way that helps match the body with sensory predictions about it (K. Friston et al. 2016). This minimization of sensory prediction error is achieved through an approximated variational Bayesian inference, with the brain storing values and generative firing rates that can be interpreted as subjective probability distributions. These variational inference mechanisms are hierarchical. Free-energy minimization and Active Inference have also been shown to connect closely with model-based reinforcement learning (K. J. Friston, Daunizeau, and Kiebel 2009; Sajid et al. 2021), a learning method actively being explored in physics-based character animation. However, the neurophysiological evidence for Free-energy minimization and Active Inference is mixed (Walsh et al. 2020; Heilbron and Chait 2018). Some critiques argue that the theory is not precise enough to be validated in scenarios like humanoid control (Kogo and Trengove 2015), or that its technical assumptions are not realistic (Raja et al. 2021; Aguilera et al. 2021). It is also difficult to grasp the specificity of Active Inference, since some of its underlying assumptions can also be found in alternative computational approaches to cognition (Sprevak 2021). What proponents and critics seem to agree on is that free-energy minimization and Active Inference have an ambition that goes far beyond task-specific motor control. In this sense, it is a general theory of cognitive development. In addition, it is an ongoing research program targeting the entirety of cognition, and therefore future work may bring additional details to it and help show more general applicability.

2.1.3.Dynamic Modelling and Joint action

A third approach to modelling human motor control is the use of dynamic modelling. Dynamic modelling uses computational models derived from non-linear physics (inertia, friction, coupled

oscillators, etc.) to model aspects of biology or psychology. For example, it has been used to model how different parts of our motor control system interact (Martin, Scholz, and Schöner 2009)
This modelling strategy has also been favoured by people studying joint action. Joint action researchers are interested in understanding motor behaviour when people engage in some joint activity (Knoblich, Butterfill, and Sebanz 2011). As such, their interest is in studying how motor coupling occurs between humans. Examples can involve spontaneous turn-taking or paying attention to the same stimuli, but the main focus is on motor synchronisation. This is considered both spontaneous (like the tendency to synchronize our steps when walking together) or voluntary (like when we try to play music together). Psychological results have shown that interpersonal synchronisation can have significant impact in terms of several aspects of prosocial behaviour (Cirelli, Einarson, and Trainor 2014; Rennung and Göritz 2016), trust, empathy or even improve learning transfer (Vink et al. 2017). In this context, quantitative models are often used as descriptive tools of how the coupling affects the behaviour of each individual (Noy, Dekel, and Alon 2011; Zhai et al. 2017; Calabrese et al. 2022). The modelling tools are derived from the physics literature, using coupled oscillators as the main modelling tool. In these scenarios, sensorimotor loops are assumed, but they remain far from being contrasted with realistic data at the perceptual or physiological level. Dynamical models are used in a more descriptive way, as opposed to optimal control theory or active inference, where models attempt to provide a functional explanation (i.e., achieve optimality, or reduce uncertainty). The mathematical assumptions are introduced to capture relations between different systems, but do not assume particular underlying principles. As such, they drive no explanation in terms of how these relations have been established, or how they affect motor learning.

Table 1: Different principle-driven theories of motor control account for some but not all of the aspects in which we are interested to understand motor control in human-character interaction scenarios

Optimal Control Theory 1. Very good account of how sensory input produces behavioural variability 2. Optimality principles help reduce the need for massive datasets 3. Not compatible with behaviour affected by emotional state 4. Doesn't consider joint action scenarios
Active Inference 1. It can explain cases where motion is only useful to improve perception 2. Potential to derive optimality principles, although cases developed are too simple 3. No focus on the extent to which behaviour is affected by emotional state, or why this happens 4. Doesn't tend to consider joint action use case
Dynamic Modelling 1. It can create rich behaviour from simple models with few parameters 2. It is difficult to address skill acquisition since it does not develop on learning mechanisms 3. Affective states can be shared or induced, but no explanation on the physiology behind it is addressed. 4. Focuses on modelling the establishment of psychological features from different kinds of behavioural and physiological coupling

2.2. Functional Anatomy of human motor control

Human motor control involves the coordinated recruitment of different parts of the nervous and muscular system. Ultimately most of the brain processing is oriented towards generating some kind of behaviour, but here we will focus on the parts that focus specifically on motor control, excluding language-related centres, sensorimotor loops, and motor planning (see Figure 1).

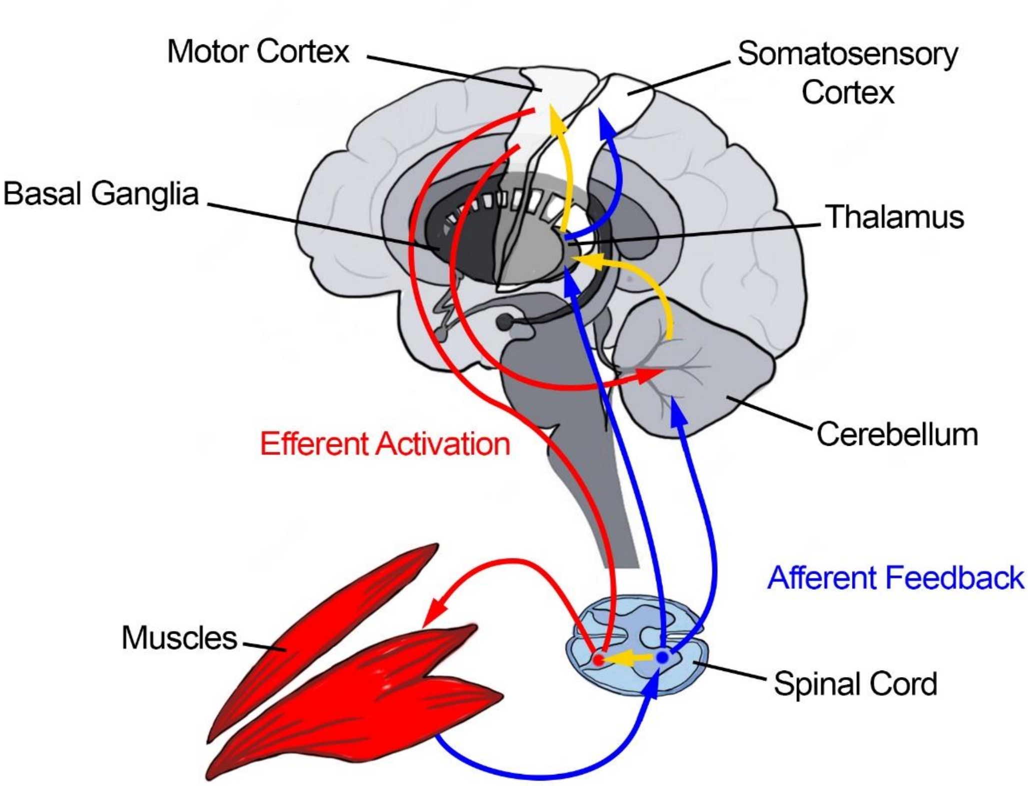

Figure 1: A schematic diagram of the anatomy of motor control and the purpose of information exchange between these areas

2.2.1.Muscles

Muscles are the actuators of our motor control system, how humans move. They generally involve pairs of them which, when activated, generate antagonistic movements (for example, an arm flexion or extension). They react to signals coming from the spinal cord (excepting the muscles related with eye and face muscles) and send proprioceptive feedback to the spinal cord from both the centre of the fibre and from the Golgi tendon organ, situated at the extreme. An important aspect of having pairs of opposed muscles is that they allow controlling not only the position and speed of the different body parts, but also the stiffness of a given articulation (Rosenbaum 2009). When used in the PBCA, these tend to be modelled as Musculo-Tendon Units (J. M. Wang et al. 2012)

2.2.2.Spinal Cord

The spinal cord acts like a transducer of the motor neurons to muscular activations. They do this within the loop that connects afferent receptors and muscle activation, and which produces the arc reflex. Therefore, the transduction of actuation signals from the brain also integrates, to some extent, the proprioception input. A significant point for later discussions on motor encoding is recurrent inhibition: Renshaw cells, which receive the signal of the motor neuron bringing the signal from the brain, can inhibit the motor signal. This is believed to help the motor signal to have more fine-grained effects. Reciprocal inhibition is also a well understood mechanism between the spine and opposing muscles, which allows regulating the stiffness of a given articulation (Rosenbaum 2009).
There is also converging evidence suggesting that the spinal cord triggers patterns of muscle activation in modules (Bizzi et al. 2008). It is also important to consider that not only actuator forces but also the stiffness of joints is actively changed in different moments of everyday tasks like walking (H. Lee,

Rouse, and Krebs 2016). Regarding connectivity, the spinal cord is essential to study sensorimotor integration. It relays sensory afferents from the muscles to the somatosensory cortex, the thalamus, and the cerebellum. It also connects the motor commands from the motor cortex with the muscles (Asan, McIntosh, and Carmel 2022).

2.2.3. Basal Ganglia

The Basal ganglia is a large brain region whose largest part is the striatum. It has motor, associative and limbic domains, and it plays a role in the execution of plans (Rosenbaum 2009), and more specifically in the performance and acquisition of new activities and tasks, the creation of habitual responses and stopping an ongoing activity to switch to a new task (Lanciego, Luquin, and Obeso 2012). This is achieved through loops that go from the brain cortex to the basal ganglia, and map back to the cortex through the thalamus (Logiaco, Abbott, and Escola 2021).

2.2.4. Motor Cortex

The motor cortex is one of the most studied parts of the human brain. Motor cortex research has traditionally differentiated between the primary motor cortex and the premotor cortex. The primary motor cortex was interpreted as triggering motor actions in different body regions. In this view different parts of the primary motor cortex controlled different muscles, and the selection of which muscles were activated was done in an earlier processing step, assumed to occur in the premotor area, following the instructions provided by other brain areas, where decisions were made in an abstract domain. In this regard, it was interpreted as a hierarchical organization, where action preparation preceded action execution. This traditional view was strongly influenced by decades-old work showing that short bursts of electrical stimuli triggered muscular twitches in different parts, and that this was more apparent in the primary motor cortex than in the premotor cortex. This vision was popularized by Penfield’s homunculus, a representation of how the different body parts were represented in the primary motor cortex (see, for example, (Rosenbaum 2009) citing the original source found in (Penfield and Rasmussen 1950)).

In recent years this view has been challenged by the work of Michael Graziano and his colleagues (M. Graziano 2006; M. S. A. Graziano and Aflalo 2007; M. S. Graziano 2016). The main claim is that the motor cortex, traditionally considered to map different body parts actually organizes behaviour in functional maps. The functional maps correspond to the most ethologically relevant behaviours (in primates). (M. Graziano 2008) highlights the very early work of Fritch and Hitchin in the late 19th century to argue these functional maps are best observed not with short stimulation bursts, but rather with stimulation that lasts at least as long as the movement targeted. In this view, the motor cortex coordinates complex actions within the behavioural repertoire. Recent evidence in the human motor cortex (Gordon et al. 2023) suggests this is a more accurate depiction of the function of the motor cortex, consistent with the idea that some aspects of action planning might derive from movement coordination.

2.2.5. Somatosensory Cortex

The somatosensory cortex is the main part of the brain cortex where proprioception signals originating in the muscles, and transmitted through the spine, are processed. It also integrates information from the visual and auditory cortices. In the same way that the primary motor cortex is associated with a map of muscles spread across the body, the somatosensory cortex is associated with somatotopy: a correspondence between body regions and parts of the somatosensory cortex (Sanchez Panchuelo et al. 2018). However, this picture has also been shown to require some nuancing. The somatosensory cortex has been shown to contain several spatial maps to control movement (M. S. Graziano and Gross 1998). In addition, spatial maps stored in the brain can represent space either in an allocentric or egocentric reference frame (Gross and Graziano 1995), and activity in the somatosensory cortex is strongly modulated by motor activity (Forss and Jousmäki 1998).

Moreover, part of the the mirror system has also been identified with the ventral part of the somatosensory cortex (Rizzolatti and Craighero 2004). The mirror system gained large and widespread interest because it showed that it responded to action representation in both the self and in others. This has shown to have considerable importance, for example, in learning to performing an action from seeing other people perform it (Ramsey, Kaplan, and Cross 2021). There are other regions involved in spatial representations related to motor planning and the mirror system, but for the purpose of this review, where we focus only on motor control, this short introduction suffices.

2.2.6. Cerebellum

The cerebellum plays a critical role in fine-grained motor control. In terms of connectivity, it receives an efferent copy of the motor cortex relayed through the brainstem, and relays back information to the cortex, via the thalamus. It also receives proprioception signals, both directly as well as mediating the thalamus (Asan, McIntosh, and Carmel 2022).

The classic model of the cerebellum, called Marr-Albus-Ito assumed it acted like a forward model (Marr 1969), and was generally considered to learn with supervised learning (Albus 1971). In this view, the cerebellum predicts an outcome from motor commands and the perceived state of the body. It is interpreted as a forward model since from a set of muscle activations and the current body pose (position or rotation of different articulations, as well as contacts) it predicts a future body pose. The efferent copy postulated by Optimal Control Theory has often been identified with the cerebellum (Miall and Wolpert 1996). The cerebellum allows the motor cortex to compare the prediction of the forward model with the perceived outcomes, as estimated in the somatosensory cortex, and adjust correspondingly the motor actions. This is why lesions in the cerebellum result in problems in fine-grained coordination. However, it seems to do more things than predicting future poses. For example, it is important to regulate muscle tone (Rosenbaum 2009), it has been proposed to play a role in intuitive reasoning (Ito 2008) and it is likely to play a role in agency (Welniarz, Worbe, and Gallea 2021). Moreover, reinforcement learning mechanisms have been proposed as a complement to supervised learning as a complementary mechanism to build the forward model of the cerebellum faster (Yamazaki and Lennon 2019).

2.2.7. Thalamus

The thalamus is a region critical to integrate different stimuli and motor actions. Thalamocortical loops are crucial to multimodal perception (for example, integrating audio and visual stimuli) and to sensorimotor integration (for example, communicating somatosensory and motor cortices). It is also critical to relay signals from the basal ganglia and from the cerebellum towards the motor cortex. As such it is a critical part of motor coordination and, in general, of sensorimotor integration (Rosenbaum 2009).

3. Tools and Strategies for physics-based character animation

3.1. Physics simulation

3.1.1. Physics Actuators and controllers

Most of recent contributions in PBCA use servo-like actuators, where torques are applied on either hinge or ball-and-socket articulations. Virtually all the controllers mentioned in this section use these. There are two reasons to choose these: each joint has only one actuator, and therefore the dimensionality of the action space is lower. This simplifies the learning of the targeted tasks. The second reason is that when we compare them servo-like actuators give more effective controllers (Peng and van de Panne 2017), possibly because they allow encoding an action space that is closer to a kinematic representation, as opposed to a force-related encoding. Since the targeted behaviours involve a spatial aspect (move forward, reach a target, etc.), this kind of encoding might be more effective.

An aspect of physics-based animation controllers which is often mentioned tangentially but is of crucial importance to reproduce the different results discussed in this section is the use of Proportional Derivative (PD) controllers. A PD controller transforms a target rotation (in angles) into a torque force to be applied. Simple, traditional PD controllers take into account the current velocity of the joint. More sophisticated PD controllers like Stable PD (Tan, Liu, and Turk 2011) or Linear Stable PD controllers (Yin and Yin 2020) also take into account the relation of weights and inertias along the hierarchy of articulated rigid bodies forming a ragdoll. Most, if not all of the contributions discussed use Stable PD or Linear Stable PD controllers. As a result, the actions to be learnt are represented as target rotations for the actuators, but indirectly integrate weights and inertias within the learning loop.

There are also some contributions that use Musculo-Tendon Units for locomotion (J. M. Wang et al. 2012) sensorimotor integration (Nakada et al. 2018) or motion retargeting (Ryu et al. 2021), but they are the minority. Most contributions in physics-based character animation use servo-like actuators.

Opposed to PBCA we will use the term **kinematic controller** to refer to methods that infer the pose of a character based on kinematic calculations, irrespective of any physical simulation. Traditionally, interactive character animation in videogames and virtual reality experiences has not been based on PBCA. Rather, it was resolved spatial constraints with methods such as direct and inverse kinematics, combined with collision detection methods.

3.1.2. Simulation engines

All PBCA controllers use a physics simulation, this is the space in which they act. Traditionally, research-oriented physics engines have been used. For example, (Peng et al. 2018; Won, Gopinath, and Hodgins 2020) used Bullet (https://pybullet.org). Works originating in the robotics literature tend to use Mujoco (https://mujoco.org/), such as for example (Merel et al. 2019). Others use DART (http://dartsim.github.io/), such as for example (Seyoung Lee et al. 2021). Finally, some of the more application-oriented work (such as (Fussell, Bergamin, and Holden 2021)) tend to use PhysX (https://developer.nvidia.com/physx-sdk), which is the de facto standard for product-oriented game engines such as Unity3D or Unreal Engine, and the ones used almost exclusively for virtual reality development. There are not many systematic comparisons between them, but (Erez, Tassa, and Todorov 2015) implemented some test comparisons and showed game engines work better to simulate games and robotic-oriented engines work better to simulate realistic constraints and collisions. It is worth noticing that each of the different engines described are software projects that progress iteratively, and therefore it may be challenging to know the extent to which the latest version of one or another software project improves performance in one or another way.

A relatively recent development, and potentially a transforming one, is the development of differentiable physics engines. The main advantage of a differentiable physics engine is that gradients of any variable involved in the physical simulation can be calculated explicitly, relative to any parameter of the simulation. Recent solutions have either been developed as novel tools (Degrave et al. 2019; de Avila Belbute-Peres et al. 2018; Freeman et al. 2021) or derived from existing physics simulators such as DART (Werling et al. 2021). As a result, PD controllers can use precise estimations of velocities in the immediate future. Crucially, deep Reinforcement Learning (RL) can use the gradient to learn much faster how a change in a force exerted on a specific rigid body will affect the physical simulation (a more detailed review of deep RL for PBCA is developed in section 3.3). Differentiable physics engines have also been combined with compilation in graphics cards (Freeman et al. 2021; Makoviychuk et al. 2021) allowing to overcome significant bottlenecks in memory management and communication between the graphics unit and the central processing unit. The consequences have been massive training speed-ups, and unlocking the creation of physics-based animation controllers using an entirely different amount of training time (see, for example (Peng et al. 2022)). It is still an open research question, though, the extent to which policies trained in one engine can transfer to another engine or to real-world robots (see, for example (Exarchos et al. 2021)).

3.2. Forward models for character control

3.2.1. Trivial examples

A passive ragdoll, as readily available in current game engines is a trivial example of a forward model for character control: the input is a configuration of articulation attached to rigid bodies with masses, plus contact forces, and the outcome is a set of rotations and velocities of the rigid bodies, calculated by the game engine.

3.2.2. Forward models for action selection

A simple example of forward model for PBCA can be found in (Grzeszczuk, Terzopoulos, and Hinton 1998). A neural network was trained to find the right actuation forces that would maximize a given objective function. For example, it would learn to find which are the inputs were needed to match a distance to a goal or minimize the deviation from a desired speed. The learning process allowed finding which are the right forces to apply to maximize the objective function.

A different use of a forward model is to model the environment. For example, in (Jordan and Rumelhart 1992) a partially learnt forward model of the environment is used to help learn the right control policy. In (L. Liu et al. 2010) it was showed that sampling in the vicinity of the current pose and running forward simulations could be used to reconstruct contact forces, typically not available in a physics engine. In both cases, the forward model helped chosing the right action to obtain a targeted movement. More recently, (Fussell, Bergamin, and Holden 2021) showed that they could use a forward model of the entire physics simulation, as a proxy of what a differentiable engine would provide: the gradient that maximizes the reward that a deep RL seeks. Since differentiable physics engines are not available in commercial game engines (in 2023 commercial VR-compatible game engines mainly use PhysX), the authors showed that this forward model could succesfully stand in as a proxy for such differentiable engine. This strategy evokes the partially learnt forward model described in (Jordan and Rumelhart 1992), and in both cases they transform the problem of training an inverse model for character control into a supervised learning problem, where the input is the current pose and the output is the targeted pose, which we can immediately calculate through the forward model. It also builds on earlier developments like (Heess et al. 2015), where it was shown that it is possible to learn to estimate the gradient.

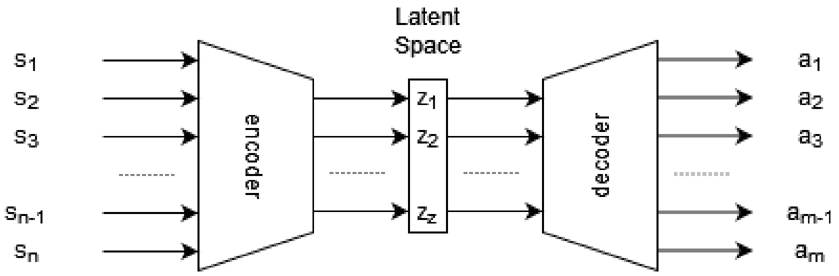

Figure 2: A diagram showing how a latent space ($z_1..z_n$) is generated by forcing dimensionality reduction between a space of states and a space of actions (i.e., $n < z$ and $m < z$)

3.2.3.Latent spaces

A machine learning construct that has shown very useful in developing PBCA controllers are variational auto-encoders (VAEs) (Kingma and Welling 2014; Rezende, Mohamed, and Wierstra 2014). An autoencoder is composed of an encoder and a decoder. The encoder is a multi-layer perceptron that encodes a large input space to a lower-dimensionality space, which is called the latent space. The decoder in turns convert the latent space to a larger space.

Due to its lower dimensionality, the latent space is generally interpreted to represent an efficient encoding of the larger dimensional input. The main difference between an autoencoder and a VAE is that a VAE allows exploring the latent space to generate a wide range of novel outputs, while maintaining the output variability within the space defined by the training data. In other terms: the space of possible outputs can be explored like a random space, but the behaviour generated will always have a resemblance with one part or another of the training dataset. An additional benefit of VAEs is that they can be trained offline, like a supervised learning problem, and therefore allow for greater generalization than RL setups.

In section 3.3.3 we will discuss how these are used in the context of inverse models: latent spaces allow training a policy that selects the appropriate actions in a latent space of actions (which has much lower dimensionality) and to evaluate the rewards in the high dimensional space. Before, we need to introduce inverse models.

3.3. Inverse models for character control

Inverse models do the opposite than forward models: given a target pose, they suggest what forces should be applied to get the target pose. To create such controllers it is useful to consider three questions:

1. How is the control of interactive behaviour approached?
2. What reference animations and other data it uses for training?
3. How is the controller trained?

Since most of recent contributions to physics-based animation controllers are based on deep RL, we will adopt its terminology to review the different answers proposed to these questions. RL is a method to train an inverse model from inputs (states) and ouptuts (actions). In recent physics-based controllers, the RL components (see **Figure 3)** are identified with physical elements as follows:

- The **environment** is any aspect of the physics simulation that affects the behaviour of the agent. This includes the body of the agent, represented as an articulated ragdoll, but also the physical simulation that calculates the dynamics update of the world in which the articulated ragdoll is embedded, the update of the rigid bodies forming the ragdoll and the actuators, which are often, but not always, Proportional Derivative (PD) controllers that map the target rotations to actual torques (and sometimes forces).
- The **states** describe the physics simulation as perceived by the agent. In physics-based animation they generally include the rigid bodies forming the articulated ragdoll that is

controlled by the physics controller (i.e., their rotation and angular velocities). Sometimes they also include the state of a reference moving ragdoll (either from reproducing an animation or controlled by a kinematic controller), parameters like the target location (for an object to grasp, or for the body to walk to), a desired speed, and sometimes also information on the terrain or any other input relevant for the task. The states are updated at each simulation step. Sometimes they are referred as perceptions, to highlight similarities with sensorimotor approaches.

- An **action** is a vector of elements to control an articulated ragdoll. These can be torques or activations of Musculo-Tendon Units for each articulated rigid body. However, they can also be target rotations or target velocities for each articulated rigid body, which are converted to torques by a proportional derivative controller.
- The **policy** is a strategy that the agent uses to decide an action, taking as basis the state of the world. It takes the form of a function mapping states to actions. Policies are learnt in a training stage, and then used in an inference stage.
- The **reward** is a scalar calculated from the states of the world. It is also updated at each simulation step. It is used in the training stage to update a value function, which is an estimate of the long-term value associated with performing an action in a given state. In physics-based character animation, training is most often done with actor-critic methods, where the policy function and the value function are trained simultaneously. The most common actor-critic are Proximal Policy Optimization (Schulman et al. 2017) or Soft actor Critic (Haarnoja et al. 2018). However, we will also see other not only actor-critic methods are used for training, in certain proposals.

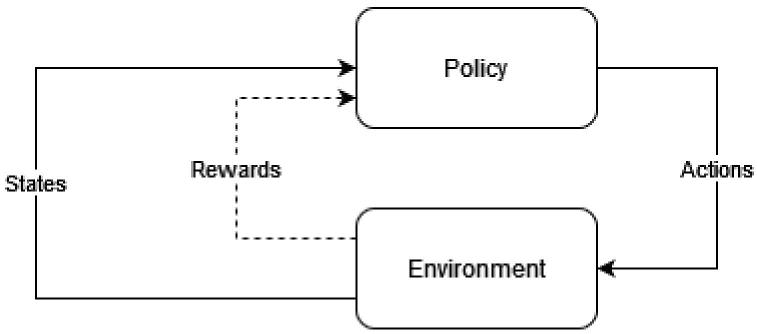

Figure 3: Main components of a RL agent, adapted from (Sutton and Barto 2018). The dashed line in rewards reflects the fact that the rewards are only used in the training stage, not at the inference stage.

The previous summary will suffice for this review. The reader interested in more details about how RL is used through the character animation literature can find a good survey in (Kwiatkowski et al. 2022).

3.3.1. Two approaches to control interactive behaviour

We now introduce a distinction in the way PBCA inverse models approach interactive control (i.e., the integration of interactive input). This distinction is not based on how training is set up, but rather on how different aspects of the environment affect the behaviour generated by the PBCA controller when it is deployed, after training (see also **Figure 4**):

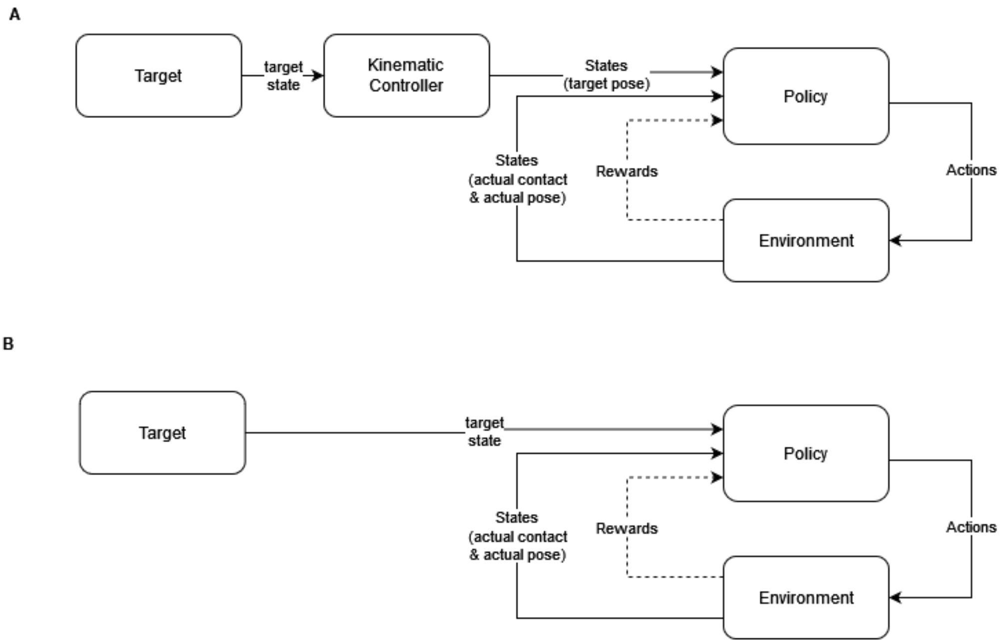

Figure 4: Two strategies for interactive control in physics-based animation controllers based on deep reinforcement learning. In both cases part of the states represent the actual pose of the character, and the contacts with the floor or other elements. A: In kinematic imitators, the target pose is determined by a kinematic controller and the task of the RL agent is to imitate it as closely as possible. B: in sensorimotor controllers the movement generated by the physics controller is more directly affected by the target states.

Kinematic imitators

Kinematic imitators focus on training a physics-based controller to imitate a kinematic controller (Bergamin et al. 2019; T. Wang et al. 2020; Won, Gopinath, and Hodgins 2020). The physics-based controller does not handle the interaction with the environment or external inputs. Rather, the interactive input is integrated with kinematic methods. For example, if we want a character to move a hand to reach an object in a particular position, or we want a character to jump upon the press of a button, a kinematic controller will generate a set of references poses, which then will be imitated by the PBCA controller. The main benefit of this approach is that kinematic controllers are a mature and well understood technology, well adapted to the needs of the videogame industry, and the physics-based controller only needs to focus on imitating a set of target poses. Kinematic controllers can also integrate some processing within the training loop. For example, (L. Liu and Hodgins 2018) integrated trajectory optimization for the kinematics step within the RL training loop used to train the RL policy for the PBCA controller. In fact, interactive kinematic controllers are also a sub-field of research in itself (Holden, Komura, and Saito 2017; Starke et al. 2019; 2019; Holden et al. 2020; Xie et al. 2022; Hong et al. 2019).

Sensorimotor controllers

Sensorimotor controllers integrate interactive input with perception-action loops. This allows us to introduce parameters that affect the motor actions generated as perceptions (Peng et al. 2018; 2021). In *sensorimotor controllers* interactivity is based on introducing target states that are included in the state of the world being trained and affect the reward. For example, if we want a character to reach an object in a particular position, the target state is having the hand at the same position as the object. If we want a character to jump, the target state is having the character reach a certain height. These target states will have different values within the training procedure, values which are often generated randomly within certain boundaries. During the inference stage the target state will take one specific value within those boundaries.

In *sensorimotor controllers* these target states directly affect the policy at the inference stage, and therefore the movement synthesized. This is opposed to *kinematic imitators*, where the part involving interaction with the environment is addressed by the kinematic controller, reducing the task of the physics-based controller to match the kinematic pose as close as possible, while preserving balance, collisions, and other physical constraints. It should be noted that this distinction is not generally discussed in the literature, we have introduced it here since each option has considerable consequences for the design of the system. For a given paper, it is generally easy to assign it to one of the two categories just by looking at the general architecture and the training dataset, as discussed next.

Despite that this distinction may seem trivial at first sight, it has considerable consequences for the controller. A summary of all the differences between *Sensorimotor controllers* and *Kinematic imitators* is found in Table 2. The following sections explain each of these aspects and detail references for each.

Table 2: Implications of different strategies for interactive motor control

	Sensorimotor controllers	**Kinematic imitators**
Purpose	Generate specific actions integrating perceptions and task-specific goals	Imitate any movement provided as input
Use of Latent Spaces	Yes, for distillation, or for acting in latent space	Yes, for acting in latent space
PD Controller	No need to feed animation pose in inference (only when training, in reward)	Feed with animation poses at every frame at inference
Metrics between animations	Only when used for reward design (mainly inspired from GANs)	Used to cluster the animations in groups
Learning architectures beyond RL	Inverse recursive control	Supervised reinforcement learning
Complementary modules	Compliance through perception of virtual displacement Time warping for movement variety	Trajectory optimization Kinematic controllers

3.3.2. Training datasets

The two different approaches to interactive control imply the use of very different datasets:

Large variability in reference motions

Large-scale motion databases are typically used as a training reference by *kinematic imitators*. To add generality to this kind of physics-based animation controllers, a large variety of movements is needed (i.e., a motion database containing different activities like walking,jumping, dancing, etc.) . A direct way to have a large variability in the reference motions is to use large amounts of animations (Won, Gopinath, and Hodgins 2020; T. Wang et al. 2020). These can be blended with different kinematic controllers. For example, (Bergamin et al. 2019) used Motion Matching (Clavet 2016; Holden et al. 2020), while (Won, Gopinath, and Hodgins 2020) used Phase-functioned Neural Networks (Holden, Komura, and Saito 2017). It is even possible to train one physics-based controller with different kinematic controllers to introduce further input variability (Won, Gopinath, and Hodgins 2020; T. Wang et al. 2020). When large motion databases are used it is important that the reference motions used in the training procedure are samples in a balanced way. This is done in order that the physics controller considers equally the imitation of the different families of reference motions.

Sometimes, instead of large motion databases, specialized kinematic controllers are used as input to the physics controller. For example, in (Hong et al. 2019; Xie et al. 2022) we find kinematic controllers specialized in football skills, which combine motion databases with sophisticated blending strategies. The output of the kinematic step is then fed to a physics controller.

Small set of (or no) reference motions

Opposed to *kinematic imitators*, *sensorimotor controllers* typically use a small set of reference motions. *Sensorimotor controllers* must modify the movement synthesized based on a target state (reach a targeted object with the hand, arrive to a targeted height with a jump). Therefore, the training must be specific for a given skill (reaching objects, jumping). This specificity comes together with a need for few animations. Each controller can either be trained from a single animation clip (Peng et al. 2018; Seyoung Lee et al. 2021) to a small number of clips (Peng et al. 2021).

This strategy has been shown to extend to the discovery of new movements without any reference motion. For example, (Yin et al. 2021) demonstrated the discovery of athletic jumping strategies and (Frezzato, Tangri, and Andrews 2022) showed realistic get-up motions could also be discovered. These two contributions have in common that they target movements which involve strict physical constraints that can guide the discovery. (Chentanez et al. 2018) a recovery agent was designed to take over the control of a physic-based character in order to bring it back close to a reference trajectory. At that stage, a second agent trained on imitating a reference animation would take over, but the recovery agent did not use the reference animations.

When no reference animation is used, constraining the style of the movement is a significant challenge. Older physics-based character animation strategies managed to learn high-dimensional control, but the style of the movement was awkward, unlike how humans move (Heess et al. 2017). (Yin et al. 2021; Frezzato, Tangri, and Andrews 2022) use a latent space of motions (see section 3.2.3) to control the style of the motion synthesised. (Tao et al. 2022) show how a physics controller learns to get up, but the insight to get natural looking motions come from using curriculum learning to constrain the controller towards using weak forces and contacts. In these cases, the physical constraints make it difficult to learn on the basis of a traditional imitation setup, but in turn facilitate the discovery of movements that look natural by reducing the space of possible actions.

3.3.3. Training policies

Model free RL

Model-Free Deep RL was the paradigm in which convincing demonstrations of example-based motor control with physics-based controllers were initially demonstrated. For example, (Heess et al. 2017) showed model-free deep RL could be used to create physics-based characters that would walk or avoid obstacles. However, in the resulting characters the movement was outlandish, it showed unnatural poses and didn't look like a human walking. This changed when (Peng et al. 2018) demonstrated how, through the use of a reference animation, an agent could learn to move forward or not fall, while preserving the animation style of the movement produced. Target movements demonstrated including walking, running, and throwing a ball, but also acrobatic movements like air kicks. For example, if the reference animation was throwing a ball, the training procedure could learn to generate different movements in order for the ball to reach different targets. This was achieved through a reward that would do a weighted sum between, on one hand, generating motions that were similar to a reference animation and, on the other hand, the achievement of goals, like reaching a target.

A major contribution of the article was recognizing the importance of combining reference state initialization and early termination. Reference state initialization reset the ragdoll in a position that was identical to the reference animation sampled in a random state. This allowed guiding the training, and was crucial to achieve complicated motions, like backflips. Early termination, in turn, ended the episode as soon as any of a set of failure conditions was met (for example, the ragdoll touching the ground with the hands, the head or the chest). At that stage the trial would be stopped and a new one would be launched. This allowed avoiding the exploration of a large space of possible movements and making the controller converge towards desirable solutions. Despite that reference state initialization and early termination were not new (see for example (Peng, Berseth, and van de Panne 2016; Heess et al. 2016)), here it was showed how it worked for a variety of movements and characters, achieving for all of them high behavioural quality.

The main limitation was that the controller would need to follow closely the animation during the training. Therefore, each controller was trained for a reference animation. A strategy to overcome this limitation was to shift the weight of the interactive control to a kinematics controller, which created *kinematic imitators*, as discussed (see also **Figure 4**).

Strategies to build beyond these limitations within the subfield of *sensorimotor controllers* were either based on the combination of use of latent spaces (see next sub-section), on the development of distance metrics between animations or on the development of co-trained modules allowing to constrain the behaviour of the physics controller (these are discussed in sub-section 4.1).

Model-based RL and latent spaces

We talk about model-based RL when we use some model of the environment to train the policy. For example, in (Heess et al. 2015) a model of the environment is trained together with the policy to estimate the gradient of the physics simulation, something that succesfully improves learning. The assumption in these cases is that the forward model will be leveraged in the training process (Levine 2018). (Merel et al. 2019) showed it was possible to encode a large amount of RL-based expert controllers into a single latent space using a two-step procedure: first, each expert controller would be trained and, second, a latent space would be used to mimic their behaviour. As a result, they showed a unique motor controller would synthesize one or another behaviour based on a parameter applied in the latent space.

Once the latent space is established, the encoder module can then be replaced with an RL controller to train other tasks. For example, (Merel et al. 2020) showed that once the latent space established, it was possible to use RL combined with curriculum learning in order to train a physics-based character to perform sequential tasks (picking up and carrying packages to a destination) and do so in a variety of spatial configurations (different spatial layouts, different object sizes). (Won, Gopinath, and Hodgins 2022) showed a VAE pre-trained merely on imitation reference motions could then be reused by replacing the encoder with a closed-loop RL agent (see **Figure 5**). The RL agent was able to learn tasks that would otherwise be challenging to learn from scratch. An important consideration was that to balance the extent to which the learning procedure would adapt the behaviour to the new rewards while maintaining the style of the reference motion, a helper module (not depicted in **Figure 5**) was introduced in parallel to the decoder and trained together with the RL policy (see also (Won, Gopinath, and Hodgins 2021)). (Peng et al. 2022) showed that within a latent space trained with reference animations allowed to discover control policies only from goal-based data, without reference animations. Recent results have also shown that such embeddings are also useful to combine policies focused on motor control with planning skills (Haotian Zhang et al. 2023; S. Liu et al. 2022). These results suggest using a latent space for motor representation can be hugely beneficial to train physics-based controllers that can be parameterized in a variety of ways.

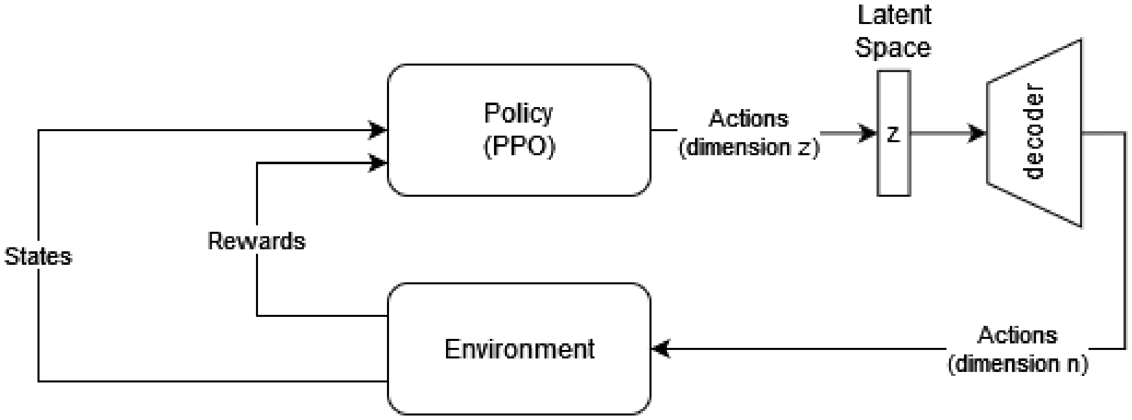

Figure 5: Once the latent space trained, the RL policy can replace the encoder, and then be trained through an output in the latent space(z < n)).

Alternatives to deep RL

All the controllers introduced so far are trained using deep-RL architectures based on actor-critic training algorithms. However, this is not a universal strategy. For example, (Fussell, Bergamin, and Holden 2021) uses supervised reinforcement learning. The article uses a large motion database for training, and the outcome can reasonably be classified as a *kinematic imitator*.

More recently, (Leibovich et al. 2022) showed it was possible to steer the learning of a dynamics controller towards a desired trajectory using iterative inversion, thus suggesting that training *sensorimotor controllers* can also be transformed into a supervised learning problem. However, this strategy has not been demonstrated with full humanoids, and it is not clear if it will scale appropriately or not.

4. The space of Motor Actions

A recurrent question commonly found in PBCA research and motor neuroscience: how are different motor actions represented in a neural controller? The contrast between PBCA and motor neuroscience is striking. We first review a variety of metrics used to compare poses and motor actions in PBCA, and then turn to motor neuroscience to discuss the extent to which PBCA can help provide a complementary angle to the question of motor encoding.

4.1. Encoding motor actions in PBCA

4.1.1. Distances between animations and latent spaces

A challenging aspect of designing a system for motor control is getting a reliable metric to measure distances between animations. In the case of physics-based animation controllers these metrics are typically used for clustering animations or for reward design.

An animation is characterized as a time series of different poses. Each pose is generally defined as a set of rotations and a root translation, where the translation is a vector 3 and the rotations are expressed either as quaternions, reduced coordinates, or matrix rotations. Comparing two animations in a simple metric is not trivial, and different metrics have been proposed for different purposes. A metric can compare different fragments of each animation irrespective of the order of each animation or compare them frame by frame respecting the temporal sequence. A metric can compare only animation chunks of the same duration or consider animation references irrespective of their duration. The metric can be defined analytically or learnt from a dataset.

When training a single controller on large sets of animations (something typical for *kinematic imitators*), it is sometimes necessary to cluster animation sequences in similar groups. This can be necessary, for example, to pick training examples evenly among the different groups, and therefore force the controller to learn evenly across different domains. In this regard, (Aristidou et al. 2018) developed a distance metric based on clustering animation databases through a deep learning classification algorithm based on a triplet loss function. However, (Won, Gopinath, and Hodgins 2020) showed a simple metric can efficiently classify a database of movement clips in clusters of similar movements. Furthermore, they showed the resulting clusters were quite useful to train a physics-based kinematic animator sampling from the different clusters in a balanced way. If we look in the field of kinematic controllers, (Starke et al. 2020) developed the notion of local motion phases to use contacts to realign movements of different durations, something which allowed introducing distances between animations per limb, instead of considering the entire pose.

When designing rewards for physics-based animation controllers, we first find analytic distances, like simple Euclidian distances between the trajectories of specific joints. For example, in (Peng et al. 2018) the quality of imitation is measured as the cumulated difference between, on one hand, the distance from the root to the end-effectors in the reference animation and, on the other hand, the same distance measured on the ragdoll controlled by the trained controller. The fact that the controller follows the animation frame by frame makes this metric very reliable. We find similar metrics in other *sensorimotor controllers*, but also in other *kinematics imitators* like (Bergamin et al. 2019). These simple metrics are used for reward quantification: the more similar the behaviour, the better the reward, and they were developed over time (see for example (L. Liu et al. 2010)).

The introduction of more subtle metrics in reward design has unlocked some of the progress in *sensorimotor controllers*. For example (Ma et al. 2021) trained physics-based animation controllers with a looser definition of space time bounds. If a controller being trained would break these space time bounds, then the trial would be early terminated (see earlier discussion of early termination in **sub-section Model free RL**, part of **section 3.3.3**) and therefore the style of the motions generated by the trained controller would respect the style defined. This allowed introducing style exploration within the behaviours synthesized by the physics controller. This style could be explored with heuristics or derived from motion datasets.

It is also worth highlighting that the use of latent spaces, as previously defined, is a way to define a metric between poses, animations or even temporal windows involving combinations of torque actuations. The latent space is where proximity is defined, and it can be explored by random exploration within the latent space. However, even very basic aspects, such as the dimensionality that we should consider for the latent space, is up for discussion. What is a good trade-off between animation fidelity and space complexity? For example, in (Won, Gopinath, and Hodgins 2022) they use a latent space of dimension 25, while in (Merel et al. 2019) they use a latent space of dimension 60. There is some work like (Nachum et al. 2019) that discusses how to measure the optimality of a latent space for motor control, but it has not been applied to humanoid control.

4.1.2. Style and affect in PBCA

More generally, the topic of Style Transfer has been quite rich in developing metrics to capture the style of a motion without capturing its actual movement. These metrics attempt to differentiate between a joyful and a sad movement, or between a tired and an aggressive movement, irresepective of whether the animation consists in walking or jumping. For example (Aberman et al. 2020) proposed a metric based on (Aristidou et al. 2018) to extract movement and style separately, using different sources. The metric was extracted through a supervised learning algorithm based on animation chunks. This allowed synthesizing a behaviour imitating the style of a given animation, although performing the movement of a second animation. (Park, Jang, and Lee 2021) used a metric developed for human action classification (Yan, Xiong, and Lin 2018). This was further expended in (Jang, Park, and Lee 2022) where it was shown that style transfer could apply to specific body parts. These works, despite interesting in the way they use distance metrics between animations, are generally based on architectures adapted from Generative Adversarial Networks (GAN) (Goodfellow et al. 2014), conditioned by the style targeted. In a GAN there is a Discriminator network, which tries to learn to distinguish between a set of reference examples and the output of a second network, the Generator. The Generator, in turn, is trained to generate images mimicking the style of a set of reference images from random noise. The criterion used for success of the Generator is that the Discriminator confounds the instances generated with the reference examples. The generative network tries to maximize the error frequency of the discriminator in a min-max game. If used directly to motion control a GAN would generate entire movement trajectories. It would therefore be impossible to use it for interactive control.

A way to introduce ideas from style transfer in *sensorimotor controllers*, while keeping their parameter-dependent nature has been through Generative Adversarial Imitation Learning (GAIL) (Ho and Ermon 2016). GAIL extends the idea of a Discriminator to interactive control by comparing expert behaviours with the ones generated by a RL agent trying to imitate the expert. Despite not being peer-reviewed, (Merel et al. 2017) and (Torabi, Warnell, and Stone 2018) are often cited as showing that GAIL could be extended to cases where the discriminator had only limited information and, when the action selected by the expert was not available, that the discriminator in a GAIL could be trained on pairs of successive states (s_t, s_{t+1}) instead of being trained on pairs of state and actions at time t (s_t, a_t). Using the metric resulting from training a discriminator from pure perception of successive states (s_t, s_{t+1}), (Xu and Karamouzas 2021) showed it was possible to train a physics-based sensorimotor controller without any reward design. Moreover, the training procedure allowed learning a graph of possible transitions, where the transitions possible were defined by the ones not triggering a "not-reference-behaviour" detection by the discriminator. Similarly, (Peng et al. 2021) showed that using a discriminator allowed training a controller to imitate not only a single reference animation, but a variety of reference movement examples. The resulting controller generated motions that looked like a variety of reference motion clips -walk, run and roll, for example-, and managed to combine these behaviours to complete a goal (reach a target with their hand, for example). In the preprint (Escontrela et al. 2022) it is even shown that these strategy performs well when transferring to real world robots.

4.1.3.Time-warping in movement dynamics

We have discussed how discriminators can be co-trained with the closed-loop Deep RL. We have also mentioned how insight into the specifics of some movements allow guiding the training procedure through reward design. For example, in (Tao et al. 2022) they constrain viable recovery movements to be weak motions, or in (Ma et al. 2021) they introduce style imitation by constraining the trajectories generated by the physics controller through space time bounds derived from animations used only as a style reference, not a movement reference.

It is also possible to use insight into the specifics of some movements to redesign the entire RL training procedure, and not only the reward aspects. This allows creating more flexible *sensorimotor controllers.* For example, (Seyoung Lee et al. 2021) modified the training setup to show time-warping could be integrated within the training procedure. They showed that parameters that affect the movement duration can also be integrated in the training. This implies it is possible to parameterize factors like the height of a jump, which not only modifies the torques applied to the articulations, but also the duration of the jump. As a result, from a single animation clip it is possible to generate an entire family of interactive movements. The approach can also be used to train an agent to push objects of variable weight. It can also be used to learn to perform the same movement with different arm lengths. This opens the door to a smart way to retarget physics-based animation controllers, transferring them between humanoids of different sizes and proportions.

4.1.4.Compliance as virtual displacements

More subtly, (Seunghwan Lee, Chang, and Lee 2022) showed that a compliance controller could provide virtual displacement information to the RL agent, who would in turn learn to minimize the contact forces between the hand and a contact object. This improved the quality of the behaviour and increased the flexibility of the controller at the same time. For example, the opening of a door would look more natural, and would also better adapt to doors of different sizes.

We have now reviewed all the relevant differences we have identified between *sensorimotor controllers* and *kinematic imitators*. For the reader interested in more detailed readings we summarize some major milestones of the last 5 years in each field in tables Table 3 and Table 4.

Table 3: References for key achievements in sensorimotor controllers between 2018 and 2023

Milestone	Reference
Flexible movement synthesis	
one controller can learn different animations	(Peng et al. 2018)
Recovery from fall	(Chentanez et al. 2018)
Discovery of plausible movements	(Yin et al. 2021)
Movement adaptation to different durations	(Seyoung Lee et al. 2021)
Compliance	(Seunghwan Lee, Chang, and Lee 2022)
Integration of objects with rich contacts	(Hassan et al. 2023)
Combination of different movements	
Hand crafted state machine	(Peng et al. 2018)
Joint policy training	(Merel et al. 2020)
Emergent from adversarial training	(Xu and Karamouzas 2021)
Emergent from parameterized latent spaces	(Peng et al. 2022)
Emergent coordination from curriculum learning	(S. Liu et al. 2022)
Combining low and high level control policies	(Haotian Zhang et al. 2023)

Table 4: References for key achievements in kinematic imitators between 2018 and 2023

Milestone	Reference
Flexible movement synthesis	
Kinematic navigation with physics-based control	(Bergamin et al. 2019)
Learning from massive motion databases	(Won, Gopinath, and Hodgins 2020)
Conversion into supervised learning problem	(Fussell, Bergamin, and Holden 2021)
Learning difficult tasks thanks to latent spaces	(Won, Gopinath, and Hodgins 2022)

4.2. Questions in human motor control

To give a richer perspective of ongoing debates in human motor control, in this section we introduce some of the recent debates in motor control, to then revisit them from the perspective of PBCA research.

4.2.1.Do motor primitives exist?

Given that PBCA mainly use PD controllers, and that PD controllers take target rotations (and velocities) as input, we might be tempted to think that, by analogy, the motor cortex represents actions in kinematic space (for example, angles and directions in space relative to a hand). This would also be useful to represent motor actions relative to surrounding objects and goals.

Alternatively, we might think that it stores motor actions as torques and combinations of torques, or an equivalent to torques that is closer to muscle activations. Intense debate on the topic and extensive experimental studies have shown that in the motor cortex we can find neurons that encode for each of those options (angle, displacement, force), and then several more cells in the motor cortex whose activity is motor related but does not fit in either (see an historical account in chapter 8 of (Lindsay 2021)). In summary, PD controllers are quite different from muscles actuators (see section 3.1.1) and we should approach any proposal regarding motor encoding with caution, since there is no reason to assume that motor encoding is not developed ad hoc, or even differently for each task.

A good way to review this cautionary tale is by revisiting the notion of motor primitives. Motor primitives have been conceptualised as sequences of actions that accomplish a goal-directed behaviour. The essential idea of a motor primitive is that there must be some "building blocks" to motor control. Motor primitives have also been identified with the underlying representation common to motor control and action recognition (Floreano, Ijspeert, and Schaal 2014). However, it has shown difficult to determine whether they should be defined in a kinematic space, or as attractors in a dynamic space of torques (Giszter 2015). Another argument for their existence has been to argue that it is impossible to learn a complex task from pure reinforcement learning, and therefore motor primitives must exist to guide this learning (Schaal and Schweighofer 2005). However, the results obtained in PBCA research directly contradict this assumption.

4.2.2. Is motor control hierarchical?

Motor primitives also help supporting the view that motor control is hierarchical. A traditional view of the motor cortex, where different parts of the primary motor cortex controlled different body parts also favored a hierarchical view. For example, (Martin, Scholz, and Schöner 2009) identify the core function of the motor cortex *to transduce the neuronal trajectory that predicts task-level motion into joint-level activation patterns.* This fits with the idea of a motor planning at a task level, using motor primitives, followed by a transduction into force-related activations. However, (Martin, Scholz, and Schöner 2009) readily admit that coupling with downstream structures may also be important to determine these activation patterns.

More recently, the view of motor control as a hierarchical system has been outlined in (Merel, Botvinick, and Wayne 2019), for both neuroscience and robotics. We agree with the authors who stated *"As research into artificial control has developed, it has become clear that in addition to task objectives, system architecture design is also critical."* In this regard there is certain similarity with PBCA, where controllers have different parts and components which, only when combined, provide interesting solutions in terms of learning efficiency, flexibility and/or quality of the behaviours synthesized. However, our previous outline on the functional anatomy of motor control (see section 2.2) suggests there is a need for significant coupling between activity in the motor and somatosensory cortices, the spine and the cerebellum. The emerging interpretation of the motor cortex we outlined suggests hierarchy within the motor cortex would be mainly ad hoc, depending on the motor movement targeted (M. Graziano 2008). Moreover, since the motor cortex also contributes to motor planning (Gordon et al. 2023), and motor control requires contributions from the somatosensory cortex and the cerebellum, it is difficult to argue that motor control is hierarchical. Opposed to this, we acknowledge that coupling between regions with different functional roles is key to generate the resulting functionality. From this perspective, it seems difficult to define cleanly the notion of motor primitives.

4.2.3. Degrees of freedom and learning priorities

A classical finding of motor neuroscience has been showing that complex patterns of movement can be generated in the spinal cord, without input from the brain. This is generally interpreted as indicating that the motor cortex does not encode completely for the action to be performed (Kalaska 2009). This is often conceptualized as the degree of freedom problem: how to define a trajectory taking into account that actuators are redundant, and that several postures allow for the same end-effector position. Moreover, several configurations of muscle activations allow for a similar pose. Other authors have argued this to be a blessing: actuator redundancy may maximize, for a given trajectory, aspects such as stability, or compliance, or other dynamic aspects that go beyond a kinematic trajectory, and such view is sometimes characterized as Synergia theory (Latash, Scholz, and Schöner 2007)

A related problem that is repeatedly mentioned in motor neuroscience is how to combine different priorities. For example, keeping balance and reaching a target with your hand. Regarding this, lambda-theory proposes that control is exerted in positional frames of reference, and that changing these is what provokes a change in posture (Feldman and Levin 1995). This idea is argued to solve several challenges, such as the problem of how motor control to reach a target does not enter in contradiction with posture-stabilizing mechanisms.(Feldman and Latash 2005).

From the perspective of PBCA controllers, combining priorities like reaching a target and keeping balance is natural, similar to any embodied agent: learning occurs in a space that requires satisfying both priorities at the same time. In other terms: the combination of different priorities seems to be naturally achieved with solutions that emerge when learning in environments that require satisfying both. There is no need for the motor controller to explicitly encode for both priorities. Similarly, having redundancy in the actuators is not a problem. Rather, it allows exploring a richer space of actions, allowing for solutions that fit better in a heterogeneous space of desired solutions combining constraints of pose, balance, etc.

4.3. A latent space in human motor encoding?

In motor neuroscience, the motor control system is characterized as a functional anatomy of heterogeneous interrelated components (cerebellum, motor cortex, spinal cord, etc.). There might not be a unique way to interpret a motor command, but there are other aspects of motor organization that seem to allow for some common ground.

If we look into recent results in the PBCA literature, a simple way to summarize all these properties might be to conceive the motor cortex being the encoder part of a VAE, and the spine the decoder part. The encoder would naturally fit the entirety of the motor repertoire in a reduced signal space, which is convenient for the anatomy of the cortico-spinal pathway, since it imposes a reduced space of signal. The VAE decoder (i.e., the spine) would be able to generate complex patterns from a reduced space of motor commands. The dimensionality reduction obtained is even more relevant if we consider actuators as MTU. MTUs are a better approximation to biological actuators than the hinge or ball-and-socket actuators associated with PD controllers, and more commonly used in PBCA research (see section 3.1.1). They also have more degrees of freedom, and therefore the problem of controlling them would benefit more from a signal transformation that involves dimensionality reduction.

In this picture the function of the motor cortex is close to Graziano’s interpretation: there is no fundamental distinction between the premotor cortex and the primary motor cortex, we do not consider action preparation associated to the premotor cortex, and specific parts of the motor cortex mapped to specific muscles. Rather, we consider the motor cortex to be encoding the repertoire of motor actions in a latent map, and integrating signals from spatial representations and motivation networks

The assumption of a latent motor space established between the motor cortex and the spine would also impose considerable constraints on the rest of the control architecture (see an outline in Figure 6). Specifically:

a. If modelled with realistic MTUs, the module standing for the function of the cerebellum would need to process information on the state of the recurrent inhibition in the different parts of the spinal cord associated with the appropriate muscles, to know how to modulate the motor commands (through the thalamus). Information regarding proprioception would also need to be processed in the cerebellum to modulate the motor commands to seek compliance with objects manipulated and other contacts. All this would probably have to occur in latent space. This is far more sophisticated than current forward models implemented in the PBCA literature.
b. The module standing for the thalamus would need to relay massively information between regions of the motor cortex, as well as the signals from the modules standing for the cerebellum and the basal ganglia towards the motor cortex. Moreover, it should do so decoding from a latent space of motor commands to an egocentric or allocentric space of motor effects. It is unclear how such connections would be established or learnt.
c. The module standing for the basal ganglia would have to handle a dual function: modulating the movement synthesized to reflect particular emotional states, and trigger switches from one action to another. However, it is unclear why or how these two functions would need to be together, motion planning and synthesis of emotional states is quite independent in humanoid motor control and PBCA.
d. It is unclear how a module standing for the somatosensory cortex would integrate proprioception signals transmitted through the spine. Proxies for the sensory streams from the visual and auditory cortices can be provided easily in a computer simulation. However, the somatosensory cortex would need to also integrate the signals coming from the proprioception input (and, eventually, decode them from the latent motor space). Physics engines typically report impulses associated with the resolution of a collision between two rigid bodies. Moreover, in physics engines the status of rigid bodies is often excluded from the simulation if they are at rest. It is therefore difficult to imagine how such data streams could be generated unless unconventional physics engines were used, or this information stream was replaced with complementary modules (for example, in a way similar to how compliance is enacted through virtual displacements in (Seunghwan Lee, Chang, and Lee 2022), similar to how virtual displacements similar

Overall, a significant effort would be needed to devise a training strategy that would, on one hand, allow controlling precisely variations in different tasks (for example, grasp an object in different positions, kick a ball in different directions) while integrating a variety of motor actions and, on the other hand, train the different modules to work consistently with each other. Most of the work cited for latent spaces is in the sub-domain of *kinematic imitators* (Won, Gopinath, and Hodgins 2022; Merel, Botvinick, and Wayne 2019) while the integration of external parameters to modify a particular behaviour is in the sub-domain of *sensorimotor controllers* (Peng et al. 2018; Seyoung Lee et al. 2021).
In summary: the notion of a latent space for motor control seems to fit naturally with different results in motor neuroscience. In PBCA research it also appears as a way to precisely control specific movements, while allowing for a variety of movements to be synthesized and combined. However, this proposal is speculative and considerable work will be required to combine all these aspects into a functioning and trainable architecture.

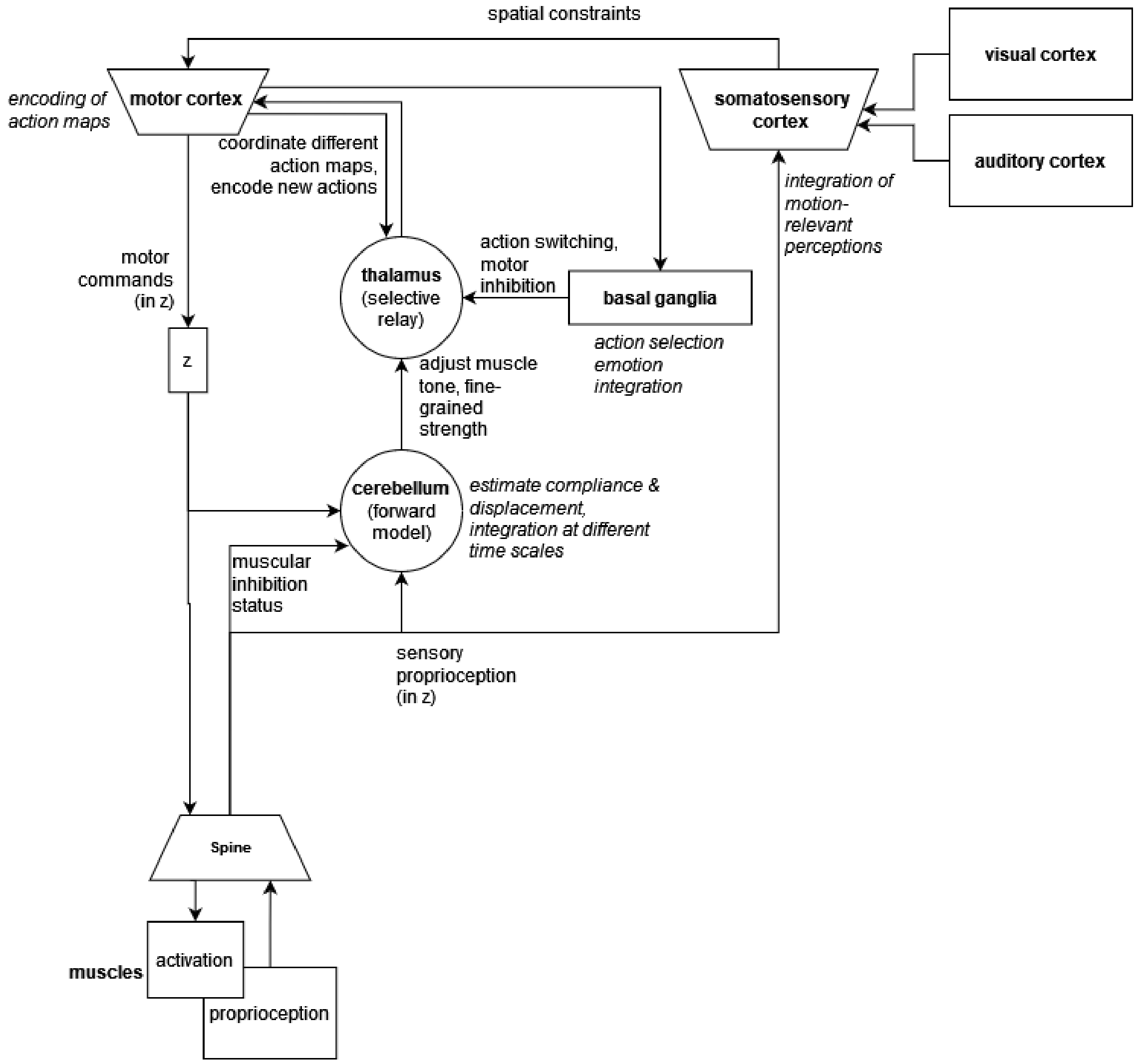

Figure 6: A schematic outline of a possible PBCA controller inspired in the functional organisation of human motor control as outlined in motor neuroscience.

5. Discussion: Physics-based animation controllers and human motor control

We have outlined how human motor control works, proposed a taxonomy of PBCA strategies and outlined a possible PBCA controller based on an interpretation of human motor control using ideas used in PBCA research. To what extent motor neuroscience and PBCA research can converge? We structure this discussion through four specific topics:

1. Sensorimotor integration
2. Skill acquisition
3. The role of emotion on behaviour control
4. Dynamic modelling and PBCA controllers

5.1. Sensorimotor integration without optimality?

Deep RL is often interpreted as being very similar to optimal control (see, for example (Merel, Botvinick, and Wayne 2019). In the context of optimal control theory, the main contribution of PBCA controllers based on Deep RL algorithms has been to show that principles similar to optimal control could be used for bodies with more degrees of freedom. This view is also compatible with how biological systems are characterized. In neuroscience, the reward is most often considered to be provided by dopamine receptors. RL models were initially validated for attentional tasks (Fiorillo, Tobler, and Schultz 2003), but dopamine receptors and the reward system is widespread enough, and it seems reasonable to assume it also affects motor control. In addition, recent developments propose that RL-based rewards are actually based on estimating probability distributions, instead of scalar values (Dabney et al. 2020). This nuances the notion of optimality associated with behaviour based on reward expectation learnt through deep RL.

However, PBCA also seem to provide a way to **test several aspects of sensorimotor integration without needing to assume optimality**, or error-minimisation. For example, it is difficult to imagine what kind of optimality criterion could be minimized to generate spontaneously synchronised behaviour similar to the one we find when people engage in joint action tasks. However, we can imagine PBCA controllers that generate such kind of coupled behaviour, and the criteria for judging their quality can be based on the quality of the behaviour synthesized when performing a task with another agent or a person, disregarding optimality in sensorimotor integration. In these scenarios the advantage of PBCA compared to models based in optimal control is that we generate full body motion. This enables performing studies similar to the ones in the joint action literature, instead of depending on constrained tasks repeated exhaustively, as is typical in behavioural studies aiming at validating sensorimotor integration mechanisms.

The main ideas behind free-energy minimization seem to be quite in contrast with droping the assumption of optimality to train PBCA controllers and mimic sensorimotor integration. A strong stance of free-energy minimization is that learning to move involves minimizing some error. Free-energy minimization and Active Inference suggest there is a unique, hierarchical, approximate, latent Bayesian model of the environment and the self, and that the overall goal of our perception and behaviour systems is to minimise a prediction error. One might argue that in most recent PBCA controllers a central part of the architecture is a reward-based agent, and that maximizing a reward is equivalent to error minimization. However, the rewards related with the outcomes of the behaviours synthesized – hitting a target, displacing forward an object, walking in a particular direction, not falling – are always related to the task being trained, and are always stated explicitly when implementing the PBCA controller. It seems unclear how task-specific rewards can be reduced to some form of general perceptual error to be minimized, as active inference seems to suggest. The motivation to perform a given action is extrinsic to the motor control needed to execute this action, at least when conceiving motor control like it is done in PBCA research. It seems difficult to consider that motor control can be driven solely by minimizing prediction error without introducing some kind of motivation aspect for action selection and planning. In this regard, it is difficult to reconcile PBCA research with free-energy minimization.

Another aspect of sensorimotor integration that is relevant for PBCA research is simulating tasks that involve rich contacts. Forward models have shown helpful to simulate tasks requiring rich contacts (L. Liu et al. 2010).Virtual displacements have also shown useful to integrate objects with compliance constraints (i.e., the opposite of stiffness) in the motion synthesized (Seunghwan Lee, Chang, and Lee 2022). More recently (Hassan et al. 2023) shown that the training strategy used in (Peng et al. 2021) can succesfully be adapted to integrate object manipulations that require rich contacts (for example, sitting in a chair, or grabbing a package and taking it somewhere). However simulations of tasks requiring sophisticated manual manipulation of objects are still based on kinematic based approaches (Chen et al. 2022; He Zhang et al. 2021), the control method does not act in the space of forces and torques. A richer representation of proprioception feedback could possibly unlock the exploration of such tasks integrating forces, and not only in the kinematic space, thus allowing more sophisticated object manipulation. However, current physics engines most often only offer contact information, instead of contact forces, and therefore pose a challenge to such a research direction.

5.2. Skill Acquisition

In motor neuroscience learning is assumed to occur through Hebbian learning. Currently we do not know through what mechanism Hebbian learning solves the credit-assignment problem. This is a key aspect of neural learning, one that in artificial neural networks is solved with back-propagation. (Payeur et al. 2021) have proposed neural burst to be this unknown mechanism and, if confirmed, this could introduce a radical change in the way we model biological networks as well as how we design artificial neural networks. An efficient implementation of Hebbian learning in artificial neural networks may bring a new generation of neuromorphic algorithms and hardware and transform entirely both computational neuroscience and deep learning engineering. Meanwhile, we cannot truly test this assumption with artificial neural networks since we have no alternative to backpropagation. We can however address skill acquisition at a slightly higher level of description.

From the perspective of PBCA, research skill transfer in (Won, Gopinath, and Hodgins 2021; 2022; Merel et al. 2020) has shown useful to address tasks that would be impossible to learn from scratch. This also occurs in humans, where skills build from the competence achieved in previously learnt tasks. In the cases where we train a controller to be able to modulate the height of a jump, the displacement of objects of different weights (Seyoung Lee et al. 2021) or simply the distance achieved when throwing a ball to a target (Peng et al. 2018), this is learnt through relearning the dynamical mapping between a temporal series of target rotations and the resulting dynamics of the movement. In the cases where there is also a compliance modulation (Seunghwan Lee, Chang, and Lee 2022), the policy learnt to implement motor control integrating compliance (i.e., minimizing contact forces) also used displacement-like inputs, even if those virtual displacements are generated in the compliance-focused module. PBCA controllers seem to provide a good opportunity to test skill acquisition hypotheses within a learning procedure based on sensorimotor integration.

Considering the space of motor commands as a latent space is also useful as a model for skill transfer. Physics-based animation controllers that use latent spaces (for example (Merel et al. 2019; Won, Gopinath, and Hodgins 2022)) suggest training in a latent space works well. (Peng et al. 2022) showed that creating a latent space trained to imitate reference motions is sufficient to discover specific skills only from goals (i.e., without further reference animations), given enough training time. This suggests the latent spaces used are an efficient encoding for motor control, in the sense that they help to reduce the dimensionality of the control problem. This can also be used to distil a unique controller for several expert behaviours or to use the trained controller into other tasks that would be difficult to learn from scratch. This occurs within a heterogeneous control network, with different parts assuming different roles: RL-based expert controllers have parameters that affect the behaviour generated (touching a target, jumping a height), while decoders translate this to the higher dimensionality of the actuators, and still PD controllers convert these signals from rotations to torques. The work developed in (Merel et al. 2019; 2020), as well as in (Peng et al. 2021) and (Won, Gopinath, and Hodgins 2022) shows a latent space established through an imitation task can be re-used to achieve a more elaborate goal that requires more complicated action coordination. This is implemented by retraining or replacing the encoder part, therefore effectively creating a better controller within the latent space. For example, (Merel et al. 2020) shows the encoder can then be retrained to sequence actions. (Peng et al. 2021) shows reaching a target can be achieved by combining more reference animations –walk, run, crouch- instead of just one reference animation like running or crouching. (Won, Gopinath, and Hodgins 2022) show that when using a latent space defined by a VAE it is possible to learn new tasks that would be very difficult to learn from scratch. (Merel et al. 2020) show that the same network used to distil different expert controllers in a common latent space can be retrained to do some task sequencing, presupposing the hierarchical nature of the task. If any of these were performed by a biological system, we would argue it reflects brain plasticity, and the capacity to adapt to novel tasks using old knowledge.

5.3. Emotion Synthesis

In section 4.1.2 we have outlined how in PBCA synthesizing a movement style that reflects a particular affective state can be trained to work independently from the task objective. In a *sensorimotor controller* the reward components that relate to movement style can be trained using reference animations and discriminators (Xu and Karamouzas 2021; Peng et al. 2021). Techniques from style transfer can also be used (Aberman et al. 2020).

In human motor neuroscience there is surprisingly little work on the impact of emotion on motor control (Rosenbaum 2009). We do know that affective valence can have an impact on movement accuracy (Coombes, Janelle, and Duley 2005). There also exists some work showing we perceive affect from motion data (Johnson, McKay, and Pollick 2011; Pollick et al. 2001). Being able to synthesize motion that shows affect independently from the motor task being performed would certainly simplify the study of affect perception. It could also help studying the impact of affective state in joint tasks. We know that factors like intimacy can affect joint motor performance (Preissmann et al. 2016), but it is challenging to study such mixed effects without being able to systematically manipulate the affective state displayed by the motion synthesized. Overall, there seem to be significant potential contributions of PBCA research in the study of the impact of emotions in motor control.

5.4. Dynamic modelling and PBCA

Approaching character control as a dynamics problem, and exploring a solution based on PBCA naturally integrates the idea of sensorimotor integration. In joint action, interpersonal coordination is assumed to emerge through the interplay of separate forward and inverse models to simulate one's own and others' actions (Keller, Novembre, and Hove 2014).

One option is to approach this problem from the perspective of *kinematic imitators*. In works like (Bergamin et al. 2019) or (Won, Gopinath, and Hodgins 2020) most of the interactivity is managed in the kinematics space, while the physical controller only integrates the outcome of the kinematic controller within the constraints of the physics simulation. Kinematics-based controllers have shown to be capable of learning physically-plausible movements like taking a seat, while respecting all the physical boundaries, including different spatial configurations and different seat sizes ((Starke et al. 2019)). It is therefore possible that they can also learn to synchronize with the behaviour of another agent, despite requiring skills to anticipate the movement of another agent.

However, we believe it has more scientific interest to explore an approach to joint action from the perspective of *sensorimotor controllers*. A key contribution to obtain a sensorimotor controller that synthesizes good quality behaviour is that in training the animation created by the controller is constantly compared with a reference animation. This reference animation (Peng et al. 2018; Seyoung Lee et al. 2021) can be interpreted as a simple and robust forward model used by the inverse controller to guide the learning process: given the current pose and forces applied, it gives an indication of which is the desired pose in the next frame, and therefore guides the search of the forces to apply.

In the brain the cerebellum acts as a forward model to ensure fine-grained coordination. It may also contribute to long-term motor memory formation of the environment in which the motor action is performed (Hikosaka et al. 2002; Imamizu et al. 2000). As previously described, the somatosensory cortex also constructs representations of the actions of others, and this has an important role in motor learning (Rizzolatti and Craighero 2004; Ramsey, Kaplan, and Cross 2021).

This suggests that building forward models that contain richer representations of reference movements and of the movements of others might contribute significantly to creating *sensorimotor controllers* for characters capable of joint synchronisation with humans in collaborative tasks. In practice, it implies modelling human sensorimotor integration as a dynamical system, as done for example in (Calabrese et al. 2022), and exploring whether integrating such computational models in sensorimotor controllers creates the desired behaviour. It therefore allows to naturally combine dynamics models of joint action with PBCA controllers. This could be a first step towards a PBCA controller closer to the outline introduced in Figure 6. The result could be a new generation of humanoid characters, capable of sophisticated interaction and coordination with humans.

Supporting sensorimotor affordances is also critical to the feeling of plausibility that can be experienced in virtual reality (Slater 2009). It therefore seems that virtual reality is a natural environment where to test such controllers in coordination with other virtual characters or humans immersed in virtual reality (Llobera et al. 2022). Virtual reality also provides precise tracking information of virtual reality users pose, gaze and even facial expressions, therefore simplifying the construction of self and other representations.

Finally, *sensorimotor controllers* also seem closer to an *enactive* perspective, as developed first in neuroscience (Varela, Thompson, and Rosch 1991) and with subsequent influence in robotics (Sandini, Metta, and Vernon 2007), where an agent learns to interact with its environment through coupled action and perception, and perception can only be understood through the idea of sensorimotor loops (Noë 2004). Enactive theory highlights the importance of motor control and sensorimotor coupling for general learning. Despite not being a theory specifically for motor control, it suggests developing such controllers might bring a new angle to the study of cognitive aspects such as agency and decision-making.

6. Conclusions

PBCA research has shown impressive developments in recent years, borrowing ideas from the larger fields of interactive character animation and machine learning. To understand how the learning modules of these systems work we have introduced a distinction between two kinds of Physics-based controllers: *sensorimotor controllers* and *kinematic animators*.

We have also reviewed the contributions of forward models of motor actions, deep RL for inverse control, the challenges involved in defining metrics between animations and how these can be used to improve PBCA controllers in different ways.

The motor neuroscience literature proposes different principled approaches to motor control, as well as specific functions for different parts of the human motor system. We have reviewed the extent to which these make sense from the perspective of PBCA controllers and proposed to interpret the combined role of the motor cortex and the spinal cord as the encoder and decoder parts of a latent space of motor actions.

We have suggested that introducing more sophisticated forward models in PBCA controllers, similarly to how the cerebellum encodes motor coordination in human motor control, may allow creating humanoid characters that collaborate significantly better with humans in a shared virtual reality.

Overall we have argued PBCA controllers can help validate empirically whether some theoretical proposals in motor neuroscience work in practice. This in turn may bring practical benefits in the fields of human-character interaction and human-robot interaction.

Appendix. Glossary of terms and acronyms used in physics-based animation

Term	Definition	First introduced
Differentiable Physics engine	A physics engine which can explicitly give a gradient of any variable involved in the physics simulation, relative to a parameter used as input	Section 3.1.2, page 12
Forward model	A computational model which, when queried with the current state of a system, answers with the predicted future state of a system.	Section 2.1.1, page 7 Section 3.2 page 13
GAIL	Generative Adversarial Imitation Learning. A training method based on comparing behaviour gneerated by an interactive controller and reference motions	Section 4.1.2, page 22
GAN	Generative Adversarial Network. A Generative method widely used in the style transfer litterature	Section 4.1.2, page 22
Interactive control	The control that is exerced integrating interactive input (for example, to hit a moving ball).	Section 3.3.1, page 15
Inverse model	A computational model which, from a desired state (typically, a pose), predicts the actions (typically, forces and torques) needed to obtain that desired state from the current state	Section 3.3., page 14
Kinematic controller	A character controller whose actions involve only positions and velocities (opposed to a Dynamics controller that also uses forces and torques in a physical simulation)	Section 3.1.1, page 12
Kinematic imitator	A physics-based animation controller that integrates interactive input with kinematic methods	Section 3.3.1 page 16
Latent space	A space of reduced dimensionality. Generally there is a mapping from the default space of perceptions or of actions through encoding and decoding modules.	Section 3.2.3, page 14
Musculo-Tendon units	Musculo-Tendon Units model pairs of opposed muscles in PBCA research. Opposed to PD controllers they allow controlling not only the position and speed of the different body parts, but also the stiffness of a given articulation	Section 2.2.1 page 9 Section 3.3.1 page 12
PBCA	Physics-Based Character Animation: A set of techniques aiming at the control of an interactive character based on applying torques and forces to different body parts	Section 1, page 5
PD	Proportional Derivative. A PD controller transforms a target rotation (in angles) into a torque force to be applied	Section 3.1.1, page 12
Policy	A strategy that the agent uses to decide an action, taking as basis the state of the world. It takes the form of a function mapping states to actions. Policies are learnt in a training stage, and then used in an inference stage.	Sectio n3.3., page 15
RL	Reinforcement Learning:A method to train an inverse model from inputs (states) and ouptuts (actions)	Section 3.3 page
Sensorimotor controller	A physics-based animation controller that integrates interactive input with perception-action loops	Section 3.3.1 page 16
VAE	Variational Auto Encoder. A method to generate a latent space that can be explored randomly to generate novel outputs.	Section 3.2.3, page 14

Acknowledgments

We would like to thank our colleagues from the Artanim Foundation: Pierre Nagorny for general feedback on machine learning and pointers on world models, Olivier Barbeau for producing the illustration in figure 1 and Bart Kevelham for suggestions on physics engines and proof-reading. We would also like to thank Balint Hodossy, PhD candidate at University College London, for pointers regarding how biological actuators are modelled and studied. We would also like to acknowledge the contribution of an anonymous reviewer for useful suggestions, constructive feedback and numerous literature pointers.

Competing interests

The authors have no interests competing with the purpose of this study.

References

Aberman, Kfir, Yijia Weng, Dani Lischinski, Daniel Cohen-Or, and Baoquan Chen. 2020. ‘Unpaired Motion Style Transfer from Video to Animation’. *ACM Transactions on Graphics (TOG)* 39 (4): 64–1.

Aguilera, Miguel, Beren Millidge, Alexander Tschantz, and Christopher L Buckley. 2021. ‘How Particular Is the Physics of the Free Energy Principle?’ *Physics of Life Reviews*.

Albus, James S. 1971. ‘A Theory of Cerebellar Function’. *Mathematical Biosciences* 10 (1): 25–61. https://doi.org/10.1016/0025-5564(71)90051-4.

Aristidou, Andreas, Daniel Cohen-Or, Jessica K Hodgins, Yiorgos Chrysanthou, and Ariel Shamir. 2018. ‘Deep Motifs and Motion Signatures’. *ACM Transactions on Graphics (TOG)* 37 (6): 1–13.

Asan, Ahmet S., James R. McIntosh, and Jason B. Carmel. 2022. ‘Targeting Sensory and Motor Integration for Recovery of Movement After CNS Injury’. *Frontiers in Neuroscience* 15. https://doi.org/10.3389/fnins.2021.791824.

Avila Belbute-Peres, Filipe de, Kevin Smith, Kelsey Allen, Josh Tenenbaum, and J. Zico Kolter. 2018. ‘End-to-End Differentiable Physics for Learning and Control’. In *Advances in Neural Information Processing Systems*, edited by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett. Vol. 31. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2018/file/842424a1d0595b76ec4fa03c46e8d755-Paper.pdf.

Beedie, Christopher J, Peter C Terry, and Andrew M Lane. 2000. ‘The Profile of Mood States and Athletic Performance: Two Meta-Analyses’. *Journal of Applied Sport Psychology* 12 (1): 49–68.

Bergamin, Kevin, Simon Clavet, Daniel Holden, and James Richard Forbes. 2019. ‘DReCon: Data-Driven Responsive Control of Physics-Based Characters’. *ACM Transactions On Graphics (TOG)* 38 (6): 1–11.

Bian, Tao, Daniel M. Wolpert, and Zhong-Ping Jiang. 2020. ‘Model-Free Robust Optimal Feedback Mechanisms of Biological Motor Control.’ *Neural Computation* 32 (3): 562–95. https://doi.org/10.1162/neco_a_01260.

Bizzi, Emilio, Vincent CK Cheung, Andrea d’Avella, Philippe Saltiel, and Matthew Tresch. 2008. ‘Combining Modules for Movement’. *Brain Research Reviews* 57 (1): 125–33. https://doi.org/10.1016/j.brainresrev.2007.08.004.

Calabrese, Carmela, Benoît G. Bardy, Pietro De Lellis, and Mario di Bernardo. 2022. ‘Modeling Frequency Reduction in Human Groups Performing a Joint Oscillatory Task’. *Frontiers in Psychology* 12. https://doi.org/10.3389/fpsyg.2021.753758.

Chen, Zoey Qiuyu, Karl Van Wyk, Yu-Wei Chao, Wei Yang, Arsalan Mousavian, Abhishek Gupta, and Dieter Fox. 2022. ‘DexTransfer: Real World Multi-Fingered Dexterous Grasping with Minimal Human Demonstrations’.

Chentanez, Nuttapong, Matthias Müller, Miles Macklin, Viktor Makoviychuk, and Stefan Jeschke. 2018. ‘Physics-Based Motion Capture Imitation with Deep Reinforcement Learning’. In *Proceedings of the 11th Annual International Conference on Motion, Interaction, and Games*, 1–10. https://doi.org/10.1145/3274247.3274506.

Cirelli, Laura K., Kathleen M. Einarson, and Laurel J. Trainor. 2014. ‘Interpersonal Synchrony Increases Prosocial Behavior in Infants’. *Developmental Science* 17 (6): 1003–11. https://doi.org/10.1111/desc.12193.

Clavet, Simon. 2016. ‘Motion Matching and the Road to Next-Gen Animation’. In *Proc. of GDC*.

Coombes, Stephen A, Christopher M Janelle, and Aaron R Duley. 2005. ‘Emotion and Motor Control: Movement Attributes Following Affective Picture Processing’. *Journal of Motor Behavior* 37 (6): 425–36.

Costa, Lancelot Da, Thomas Parr, Noor Sajid, Sebastijan Veselic, Victorita Neacsu, and Karl Friston. 2020. ‘Active Inference on Discrete State-Spaces: A Synthesis’. *Journal of Mathematical Psychology* 99: 102447. https://doi.org/10.1016/j.jmp.2020.102447.

Dabney, Will, Zeb Kurth-Nelson, Naoshige Uchida, Clara Kwon Starkweather, Demis Hassabis, Rémi Munos, and Matthew Botvinick. 2020. ‘A Distributional Code for Value in Dopamine-Based Reinforcement Learning’. *Nature* 577 (7792): 671–75. https://doi.org/10.1038/s41586-019-1924-6.

Degrave, Jonas, Michiel Hermans, Joni Dambre, and others. 2019. ‘A Differentiable Physics Engine for Deep Learning in Robotics’. *Frontiers in Neurorobotics*, 6. https://doi.org/10.3389/fnbot.2019.00006.

Erez, Tom, Yuval Tassa, and Emanuel Todorov. 2015. ‘Simulation Tools for Model-Based Robotics: Comparison of Bullet, Havok, Mujoco, Ode and Physx’. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 4397–4404. IEEE. https://doi.org/10.1109/ICRA.2015.7139807.

Escontrela, Alejandro, Xue Bin Peng, Wenhao Yu, Tingnan Zhang, Atil Iscen, Ken Goldberg, and Pieter Abbeel. 2022. ‘Adversarial Motion Priors Make Good Substitutes for Complex Reward Functions’. arXiv. https://doi.org/10.48550/ARXIV.2203.15103.

Exarchos, Ioannis, Yifeng Jiang, Wenhao Yu, and C. Karen Liu. 2021. ‘Policy Transfer via Kinematic Domain Randomization and Adaptation’. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 45–51. https://doi.org/10.1109/ICRA48506.2021.9561982.

Feldman, Anatol G, and Mark L Latash. 2005. ‘Testing Hypotheses and the Advancement of Science: Recent Attempts to Falsify the Equilibrium Point Hypothesis’. *Experimental Brain Research* 161: 91–103.

Feldman, Anatol G, and Mindy F Levin. 1995. ‘The Origin and Use of Positional Frames of Reference in Motor Control’. *Behavioral and Brain Sciences* 18 (4): 723–44.

Fiorillo, Christopher D, Philippe N Tobler, and Wolfram Schultz. 2003. ‘Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons’. *Science* 299 (5614): 1898–1902. https://doi.org/10.1126/science.1077349.

Floreano, Dario, Auke Jan Ijspeert, and Stefan Schaal. 2014. ‘Robotics and Neuroscience’. *Current Biology* 24 (18): R910–20. https://doi.org/10.1016/j.cub.2014.07.058.

Forss, Nina, and Veikko Jousmäki. 1998. ‘Sensorimotor Integration in Human Primary and Secondary Somatosensory Cortices’. *Brain Research* 781 (1–2): 259–67.

Freeman, C. Daniel, Erik Frey, Anton Raichuk, Sertan Girgin, Igor Mordatch, and Olivier Bachem. 2021. ‘Brax - A Differentiable Physics Engine for Large Scale Rigid Body Simulation’. In *Thirty-Fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 1)*. https://github.com/google/brax.

Frezzato, A, A Tangri, and S Andrews. 2022. ‘Synthesizing Get-Up Motions for Physics-Based Characters’. *Eurographics Symposium on Computer Animation* 41 (8).

Friston, Karl. 2009. ‘The Free-Energy Principle: A Rough Guide to the Brain?’ *Trends in Cognitive Sciences* 13 (7): 293–301. https://doi.org/10.1016/j.tics.2009.04.005.

———. 2010. ‘The Free-Energy Principle: A Unified Brain Theory?’ *Nature Reviews Neuroscience* 11 (2): 127–38. https://doi.org/10.1038/nrn2787.

Friston, Karl, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, John O’Doherty, and Giovanni Pezzulo. 2016. ‘Active Inference and Learning’. *Neuroscience & Biobehavioral Reviews* 68: 862–79. https://doi.org/10.1016/j.neubiorev.2016.06.022.

Friston, Karl J, Jean Daunizeau, and Stefan J Kiebel. 2009. ‘Reinforcement Learning or Active Inference?’ *PloS One* 4 (7): e6421.

Fussell, Levi, Kevin Bergamin, and Daniel Holden. 2021. ‘SuperTrack: Motion Tracking for Physically Simulated Characters Using Supervised Learning’. *ACM Transactions on Graphics (TOG)* 40 (6): 1–13.

Giszter, Simon F. 2015. ‘Motor Primitives—New Data and Future Questions’. *Current Opinion in Neurobiology* 33: 156–65. https://doi.org/10.1016/j.conb.2015.04.004.

Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. ‘Generative Adversarial Nets’. In *Advances in Neural Information Processing Systems*, 2672–80.

Gordon, Evan M., Roselyne J. Chauvin, Andrew N. Van, Aishwarya Rajesh, Ashley Nielsen, Dillan J. Newbold, Charles J. Lynch, et al. 2023. ‘A Somato-Cognitive Action Network Alternates with Effector Regions in Motor Cortex’. *Nature*, April. https://doi.org/10.1038/s41586-023-05964-2.

Graziano, Michael. 2006. ‘The Organization of Behavioral Repertoire in Motor Cortex’. *Annual Review of Neuroscience* 29: 105—134. https://doi.org/10.1146/annurev.neuro.29.051605.112924.

———. 2008. *The Intelligent Movement Machine: An Ethological Perspective on the Primate Motor System*. Oxford University Press.

Graziano, Michael S. A., and Tyson N. Aflalo. 2007. ‘Mapping Behavioral Repertoire onto the Cortex’. *Neuron* 56 (2): 239–51. https://doi.org/10.1016/j.neuron.2007.09.013.

Graziano, Michael SA. 2016. ‘Ethological Action Maps: A Paradigm Shift for the Motor Cortex’. *Trends in Cognitive Sciences* 20 (2): 121–32.

Graziano, Michael SA, and Charles G Gross. 1998. ‘Spatial Maps for the Control of Movement’. *Current Opinion in Neurobiology* 8 (2): 195–201.

Gross, Charles G, and Michael SA Graziano. 1995. ‘Review■: Multiple Representations of Space in the Brain’. *The Neuroscientist* 1 (1): 43–50.

Grzeszczuk, Radek, Demetri Terzopoulos, and Geoffrey Hinton. 1998. ‘Neuroanimator: Fast Neural Network Emulation and Control of Physics-Based Models’. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, 9–20.

Haarnoja, Tuomas, Aurick Zhou, Pieter Abbeel, and Sergey Levine. 2018. ‘Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor’. In *Proceedings of the 35th International Conference on Machine Learning*, edited by Jennifer Dy and Andreas Krause, 80:1861–70. Proceedings of Machine Learning Research. PMLR. https://proceedings.mlr.press/v80/haarnoja18b.html.

Hassan, Mohamed, Yunrong Guo, Tingwu Wang, Michael Black, Sanja Fidler, and Xue Bin Peng. 2023. ‘Synthesizing Physical Character-Scene Interactions’.

Hausmann, Sebastien B., Alessandro Marin Vargas, Alexander Mathis, and Mackenzie W. Mathis. 2021. ‘Measuring and Modeling the Motor System with Machine Learning’. *Current Opinion in Neurobiology* 70: 11–23. https://doi.org/10.1016/j.conb.2021.04.004.

Heess, Nicolas, Srinivasan Sriram, Jay Lemmon, Josh Merel, Greg Wayne, Yuval Tassa, Tom Erez, et al. 2017. ‘Emergence of Locomotion Behaviours in Rich Environments’. *ArXiv Preprint ArXiv:1707.02286*.

Heess, Nicolas, Gregory Wayne, David Silver, Timothy Lillicrap, Tom Erez, and Yuval Tassa. 2015. ‘Learning Continuous Control Policies by Stochastic Value Gradients’. In *Advances in Neural Information Processing Systems*, edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett. Vol. 28. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2015/file/148510031349642de5ca0c544f31b2ef-Paper.pdf.

Heess, Nicolas, Gregory Wayne, Yuval Tassa, Timothy P. Lillicrap, Martin A. Riedmiller, and David Silver. 2016. ‘Learning and Transfer of Modulated Locomotor Controllers’. *CoRR* abs/1610.05182. http://arxiv.org/abs/1610.05182.

Heilbron, Micha, and Maria Chait. 2018. ‘Great Expectations: Is There Evidence for Predictive Coding in Auditory Cortex?’ *Neuroscience* 389: 54–73. https://doi.org/10.1016/j.neuroscience.2017.07.061.

Hikosaka, Okihide, Kae Nakamura, Katsuyuki Sakai, and Hiroyuki Nakahara. 2002. ‘Central Mechanisms of Motor Skill Learning’. *Current Opinion in Neurobiology* 12 (2): 217–22. https://doi.org/10.1016/S0959-4388(02)00307-0.

Ho, Jonathan, and Stefano Ermon. 2016. ‘Generative Adversarial Imitation Learning’. In *Advances in Neural Information Processing Systems*, edited by D. Lee, M.

Sugiyama, U. Luxburg, I. Guyon, and R. Garnett. Vol. 29. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2016/file/cc7e2b878868cbae992d1fb743995d8f-Paper.pdf.

Holden, Daniel, Oussama Kanoun, Maksym Perepichka, and Tiberiu Popa. 2020. 'Learned Motion Matching'. *ACM Transactions on Graphics (TOG)* 39 (4): 53–1.

Holden, Daniel, Taku Komura, and Jun Saito. 2017. 'Phase-Functioned Neural Networks for Character Control'. *ACM Transactions on Graphics (TOG)* 36 (4): 42.

Hong, Seokpyo, Daseong Han, Kyungmin Cho, Joseph S Shin, and Junyong Noh. 2019. 'Physics-Based Full-Body Soccer Motion Control for Dribbling and Shooting'. *ACM Transactions on Graphics (TOG)* 38 (4): 1–12. https://doi.org/10.1145/3306346.3322963.

Imamizu, Hiroshi, Satoru Miyauchi, Tomoe Tamada, Yuka Sasaki, Ryousuke Takino, Benno Pütz, Toshinori Yoshioka, and Mitsuo Kawato. 2000. 'Human Cerebellar Activity Reflecting an Acquired Internal Model of a New Tool'. *Nature* 403 (6766): 192–95. https://doi.org/10.1038/35003194.

Ito, Masao. 2008. 'Control of Mental Activities by Internal Models in the Cerebellum'. *Nature Reviews Neuroscience* 9 (4): 304–13. https://doi.org/10.1038/nrn2332.

Jang, Deok-Kyeong, Soomin Park, and Sung-Hee Lee. 2022. 'Motion Puzzle: Arbitrary Motion Style Transfer by Body Part'. *ACM Transactions on Graphics (TOG)*. https://doi.org/10.1145/3516429.

Johnson, Kerri L., Lawrie S. McKay, and Frank E. Pollick. 2011. 'He Throws like a Girl (but Only When He's Sad): Emotion Affects Sex-Decoding of Biological Motion Displays'. *Cognition* 119 (2): 265–80. https://doi.org/10.1016/j.cognition.2011.01.016.

Jordan, Michael I, and David E Rumelhart. 1992. 'Forward Models: Supervised Learning with a Distal Teacher'. *Cognitive Science* 16 (3): 307–54.

Kalaska, John F. 2009. 'From Intention to Action: Motor Cortex and the Control of Reaching Movements'. *Progress in Motor Control: A Multidisciplinary Perspective*, 139–78.

Keller, Peter E, Giacomo Novembre, and Michael J Hove. 2014. 'Rhythm in Joint Action: Psychological and Neurophysiological Mechanisms for Real-Time Interpersonal Coordination'. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369 (1658): 20130394. https://doi.org/10.1098/rstb.2013.0394.

Kingma, Diederik P., and Max Welling. 2014. 'Auto-Encoding Variational Bayes'. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, edited by Yoshua Bengio and Yann LeCun. http://arxiv.org/abs/1312.6114.

Knoblich, Günther, Stephen Butterfill, and Natalie Sebanz. 2011. 'Psychological Research on Joint Action: Theory and Data'. In , 54:59–101. The psychology of learning and Motivation: Elsevier.

Kogo, N., and C. Trengove. 2015. 'Is Predictive Coding Theory Articulated Enough to Be Testable?' *Frontiers in Computational Neuroscience* 9.

Kwiatkowski, Ariel, Eduardo Alvarado, Vicky Kalogeiton, C Karen Liu, Julien Pettré, Michiel van de Panne, and Marie-Paule Cani. 2022. 'A Survey on Reinforcement Learning Methods in Character Animation'. In *Computer Graphics Forum*, 41:613–39. Wiley Online Library. https://doi.org/10.1111/cgf.14504.

La Barbera, Vittorio, Fabio Pardo, Yuval Tassa, Monica Daley, Christopher Richards, Petar Kormushev, and John Hutchinson. 2021. 'OstrichRL: A Musculoskeletal Ostrich Simulation to Study Bio-Mechanical Locomotion'. In *Deep RL Workshop NeurIPS 2021*.

Lanciego, José L., Natasha Luquin, and José A. Obeso. 2012. 'Functional Neuroanatomy of the Basal Ganglia.' *Cold Spring Harbor Perspectives in Medicine* 2 (12): a009621. https://doi.org/10.1101/cshperspect.a009621.

Latash, Mark L., John P. Scholz, and Gregor Schöner. 2007. 'Toward a New Theory of Motor Synergies'. *Motor Control* 11 (3): 276–308. https://doi.org/10.1123/mcj.11.3.276.

Lee, Hyunglae, Elliott J. Rouse, and Hermano Igo Krebs. 2016. ‘Summary of Human Ankle Mechanical Impedance During Walking’. *IEEE Journal of Translational Engineering in Health and Medicine* 4: 1–7. https://doi.org/10.1109/JTEHM.2016.2601613.

Lee, Seunghwan, Phil Sik Chang, and Jehee Lee. 2022. ‘Deep Compliant Control’. In *ACM SIGGRAPH 2022 Conference Proceedings*. SIGGRAPH ’22. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3528233.3530719.

Lee, Seyoung, Sunmin Lee, Yongwoo Lee, and Jehee Lee. 2021. ‘Learning a Family of Motor Skills from a Single Motion Clip’. *ACM Transactions on Graphics (TOG)* 40 (4): 1–13.

Leibovich, Gal, Guy Jacob, Or Avner, Gal Novik, and Aviv Tamar. 2022. ‘Learning Control by Iterative Inversion’. In *NeurIPS 2022 Foundation Models for Decision Making Workshop*. https://openreview.net/forum?id=pFCJiH-tAyj.

Levine, Sergey. 2018. ‘Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review’. *CoRR* abs/1805.00909. http://arxiv.org/abs/1805.00909.

Lillicrap, Timothy P., and Stephen H. Scott. 2013. ‘Preference Distributions of Primary Motor Cortex Neurons Reflect Control Solutions Optimized for Limb Biomechanics’. *Neuron* 77 (1): 168–79. https://doi.org/10.1016/j.neuron.2012.10.041.

Lindsay, Grace. 2021. *Models of the Mind: How Physics, Engineering and Mathematics Have Shaped Our Understanding of the Brain*. Bloomsbury Publishing.

Liu, Libin, and Jessica Hodgins. 2018. ‘Learning Basketball Dribbling Skills Using Trajectory Optimization and Deep Reinforcement Learning’. *ACM Transactions on Graphics (TOG)* 37 (4): 1–14. https://doi.org/10.1145/3197517.3201315.

Liu, Libin, KangKang Yin, Michiel van de Panne, Tianjia Shao, and Weiwei Xu. 2010. ‘Sampling-Based Contact-Rich Motion Control’. *ACM Trans. Graph.* 29 (4): 128:1-128:10. https://doi.org/10.1145/1778765.1778865.

Liu, Siqi, Guy Lever, Zhe Wang, Josh Merel, S. M. Ali Eslami, Daniel Hennes, Wojciech M. Czarnecki, et al. 2022. ‘From Motor Control to Team Play in Simulated Humanoid Football’. *Science Robotics* 7 (69): eabo0235. https://doi.org/10.1126/scirobotics.abo0235.

Llobera, Joan, Valentin Jacquat, Carmela Calabrese, and Caecilia Charbonnier. 2022. ‘Playing the Mirror Game in Virtual Reality with an Autonomous Character’. *Scientific Reports*, no. 12, 21329. https://doi.org/10.1038/s41598-022-25197-z.

Logiaco, Laureline, L. F. Abbott, and Sean Escola. 2021. ‘Thalamic Control of Cortical Dynamics in a Model of Flexible Motor Sequencing.’ *Cell Reports* 35 (9): 109090. https://doi.org/10.1016/j.celrep.2021.109090.

Ma, Li-Ke, Zeshi Yang, Xin Tong, Baining Guo, and KangKang Yin. 2021. ‘Learning and Exploring Motor Skills with Spacetime Bounds’. In *Computer Graphics Forum*, 40:251–63. Wiley Online Library. https://doi.org/10.1111/cgf.142630.

Makoviychuk, Viktor, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, et al. 2021. ‘Isaac Gym: High Performance GPU Based Physics Simulation For Robot Learning’. In *Thirty-Fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*. https://openreview.net/forum?id=fgFBtYgJQX_.

Marr, David. 1969. ‘A Theory of Cerebellar Cortex’. *The Journal of Physiology* 202 (2): 437–70. https://doi.org/10.1113/jphysiol.1969.sp008820.

Martin, V., J. P. Scholz, and G. Schöner. 2009. ‘Redundancy, Self-Motion, and Motor Control’. *Neural Computation* 21 (5): 1371–1414. https://doi.org/10.1162/neco.2008.01-08-698.

Mathis, Mackenzie W., and Steffen Schneider. 2021. ‘Motor Control: Neural Correlates of Optimal Feedback Control Theory’. *Current Biology* 31 (7): R356–58. https://doi.org/10.1016/j.cub.2021.01.087.

Merel, Josh, Matthew Botvinick, and Greg Wayne. 2019. ‘Hierarchical Motor Control in Mammals and Machines’. *Nature Communications* 10 (1): 1–12. https://doi.org/10.1038/s41467-019-13239-6.

Merel, Josh, Leonard Hasenclever, Alexandre Galashov, Arun Ahuja, Vu Pham, Greg Wayne, Yee Whye Teh, and Nicolas Heess. 2019. ‘Neural Probabilistic Motor Primitives for Humanoid Control’. In *International Conference on Learning Representations*.

Merel, Josh, Yuval Tassa, Sriram Srinivasan, Jay Lemmon, Ziyu Wang, Greg Wayne, and Nicolas Heess. 2017. ‘Learning Human Behaviors from Motion Capture by Adversarial Imitation’. *ArXiv Preprint ArXiv:1707.02201*.

Merel, Josh, Saran Tunyasuvunakool, Arun Ahuja, Yuval Tassa, Leonard Hasenclever, Vu Pham, Tom Erez, Greg Wayne, and Nicolas Heess. 2020. ‘Catch & Carry: Reusable Neural Controllers for Vision-Guided Whole-Body Tasks’. *ACM Transactions on Graphics (TOG)* 39 (4): 39–1.

Miall, R Chris, and Daniel M Wolpert. 1996. ‘Forward Models for Physiological Motor Control’. *Neural Networks* 9 (8): 1265–79. https://doi.org/10.1016/S0893-6080(96)00035-4.

Nachum, Ofir, Shixiang Gu, Honglak Lee, and Sergey Levine. 2019. ‘Near-Optimal Representation Learning for Hierarchical Reinforcement Learning’. In *International Conference on Learning Representations*. https://openreview.net/forum?id=H1emus0qF7.

Nakada, Masaki, Tao Zhou, Honglin Chen, Tomer Weiss, and Demetri Terzopoulos. 2018. ‘Deep Learning of Biomimetic Sensorimotor Control for Biomechanical Human Animation’. *ACM Transactions on Graphics (TOG)* 37 (4): 56.

Noë, Alva. 2004. *Action in Perception*. MIT press.

Noy, Lior, Erez Dekel, and Uri Alon. 2011. ‘The Mirror Game as a Paradigm for Studying the Dynamics of Two People Improvising Motion Together’. *Proceedings of the National Academy of Sciences of the United States of America* 108 (52): 20947–52. https://doi.org/10.1073/pnas.1108155108.

Park, Soomin, Deok-Kyeong Jang, and Sung-Hee Lee. 2021. ‘Diverse Motion Stylization for Multiple Style Domains via Spatial-Temporal Graph-Based Generative Model’. *Proceedings of the ACM on Computer Graphics and Interactive Techniques* 4 (3): 1–17. https://doi.org/10.1145/3480145.

Payeur, Alexandre, Jordan Guerguiev, Friedemann Zenke, Blake A Richards, and Richard Naud. 2021. ‘Burst-Dependent Synaptic Plasticity Can Coordinate Learning in Hierarchical Circuits’. *Nature Neuroscience* 24 (7): 1010–19. https://doi.org/10.1038/s41593-021-00857-x.

Penfield, Wilder, and Theodore Rasmussen. 1950. *The Cerebral Cortex of Man; a Clinical Study of Localization of Function.* Macmillan.

Peng, Xue Bin, Pieter Abbeel, Sergey Levine, and Michiel van de Panne. 2018. ‘DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills’. *ACM Transactions on Graphics (TOG)* 37 (4): 143:1-143:14. https://doi.org/10.1145/3197517.3201311.

Peng, Xue Bin, Glen Berseth, and Michiel van de Panne. 2016. ‘Terrain-Adaptive Locomotion Skills Using Deep Reinforcement Learning’. *ACM Trans. Graph.* 35 (4). https://doi.org/10.1145/2897824.2925881.

Peng, Xue Bin, Yunrong Guo, Lina Halper, Sergey Levine, and Sanja Fidler. 2022. ‘ASE: Large-Scale Reusable Adversarial Skill Embeddings for Physically Simulated Characters’. *ACM Trans. Graph.* 41 (4).

Peng, Xue Bin, Ze Ma, Pieter Abbeel, Sergey Levine, and Angjoo Kanazawa. 2021. ‘AMP: Adversarial Motion Priors for Stylized Physics-Based Character Control’. *ACM Trans. Graph.* 40 (4). https://doi.org/10.1145/3450626.3459670.

Peng, Xue Bin, and Michiel van de Panne. 2017. ‘Learning Locomotion Skills Using DeepRL: Does the Choice of Action Space Matter?’ In *Proceedings of the ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, 12:1-12:13. SCA ’17. New York, NY, USA: ACM. https://doi.org/10.1145/3099564.3099567.

Pollick, Frank E., Helena M. Paterson, Armin Bruderlin, and Anthony J. Sanford. 2001. ‘Perceiving Affect from Arm Movement’. *Cognition* 82 (2): B51–61. https://doi.org/10.1016/S0010-0277(01)00147-0.

Preissmann, Delphine, Caecilia Charbonnier, Sylvain Chagué, Jean-Philippe Antonietti, Joan Llobera, Francois Ansermet, and Pierre J Magistretti. 2016. ‘A Motion Capture Study to Measure the Feeling of Synchrony in Romantic Couples and in Professional Musicians’. *Frontiers in Psychology* 7: 1673.

Raja, Vicente, Dinesh Valluri, Edward Baggs, Anthony Chemero, and Michael L. Anderson. 2021. ‘The Markov Blanket Trick: On the Scope of the Free Energy Principle and Active Inference’. *Physics of Life Reviews*. https://doi.org/10.1016/j.plrev.2021.09.001.

Ramsey, Richard, David M Kaplan, and Emily S Cross. 2021. ‘Watch and Learn: The Cognitive Neuroscience of Learning from Others’ Actions’. *Trends in Neurosciences* 44 (6): 478–91.

Rennung, Miriam, and Anja S. Göritz. 2016. ‘Prosocial Consequences of Interpersonal Synchrony’. *Zeitschrift Für Psychologie* 224 (3): 168–89. https://doi.org/10.1027/2151-2604/a000252.

Rezende, Danilo Jimenez, Shakir Mohamed, and Daan Wierstra. 2014. ‘Stochastic Backpropagation and Approximate Inference in Deep Generative Models’. In *International Conference on Machine Learning*, 1278–86. PMLR.

Rizzolatti, Giacomo, and Laila Craighero. 2004. ‘The Mirror-Neuron System’. *Annu. Rev. Neurosci.* 27: 169–92.

Rosenbaum, David A. 2009. *Human Motor Control*. Academic press.

Ryu, Hoseok, Minseok Kim, Seungwhan Lee, Moon Seok Park, Kyoungmin Lee, and Jehee Lee. 2021. ‘Functionality-Driven Musculature Retargeting’. In *Computer Graphics Forum*, 40:341–56. Wiley Online Library.

Sajid, Noor, Philip J. Ball, Thomas Parr, and Karl J. Friston. 2021. ‘Active Inference: Demystified and Compared’. *Neural Computation* 33 (3): 674–712. https://doi.org/10.1162/neco_a_01357.

Sanchez Panchuelo, Rosa M., Julien Besle, Denis Schluppeck, Miles Humberstone, and Susan Francis. 2018. ‘Somatotopy in the Human Somatosensory System’. *Frontiers in Human Neuroscience* 12. https://doi.org/10.3389/fnhum.2018.00235.

Sandini, Giulio, Giorgio Metta, and David Vernon. 2007. ‘The ICub Cognitive Humanoid Robot: An Open-System Research Platform for Enactive Cognition’. In *50 Years of Artificial Intelligence: Essays Dedicated to the 50th Anniversary of Artificial Intelligence*, edited by Max Lungarella, Fumiya Iida, Josh Bongard, and Rolf Pfeifer, 358–69. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77296-5_32.

Schaal, Stefan, and Nicolas Schweighofer. 2005. ‘Computational Motor Control in Humans and Robots’. *Current Opinion in Neurobiology* 15 (6): 675–82. https://doi.org/10.1016/j.conb.2005.10.009.

Schulman, John, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. ‘Proximal Policy Optimization Algorithms’. *CoRR* abs/1707.06347. http://arxiv.org/abs/1707.06347.

Scott, Stephen H. 2004. ‘Optimal Feedback Control and the Neural Basis of Volitional Motor Control’. *Nature Reviews Neuroscience* 5 (7): 532–45. https://doi.org/10.1038/nrn1427.

Slater, M. 2009. ‘Place Illusion and Plausibility in Virtual Environments’. *Philos Trans R Soc Lond*.

Sprevak, Mark. 2021. ‘Predictive Coding I: Introduction’. http://philsci-archive.pitt.edu/19365/.

Starke, Sebastian, He Zhang, Taku Komura, and Jun Saito. 2019. ‘Neural State Machine for Character-Scene Interactions.’ *ACM Trans. Graph.* 38 (6): 209–1.

Starke, Sebastian, Yiwei Zhao, Taku Komura, and Kazi Zaman. 2020. ‘Local Motion Phases for Learning Multi-Contact Character Movements’. *ACM Transactions on Graphics (TOG)* 39 (4): 54–1.

Sutton, Richard S, and Andrew G Barto. 2018. *Reinforcement Learning: An Introduction*. MIT press.

Tan, Jie, Karen Liu, and Greg Turk. 2011. 'Stable Proportional-Derivative Controllers'. *IEEE Computer Graphics and Applications* 31 (4): 34–44.

Tao, Tianxin, Matthew Wilson, Ruiyu Gou, and Michiel van de Panne. 2022. 'Learning to Get Up'. In *ACM SIGGRAPH 2022 Conference Proceedings*. SIGGRAPH '22. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3528233.3530697.

Todorov, Emanuel. 2004. 'Optimality Principles in Sensorimotor Control'. *Nature Neuroscience* 7 (9): 907–15. https://doi.org/10.1038/nn1309.

Todorov, Emanuel, and Michael I Jordan. 2002. 'Optimal Feedback Control as a Theory of Motor Coordination'. *Nature Neuroscience* 5 (11): 1226–35. https://doi.org/10.1038/nn963.

Torabi, Faraz, Garrett Warnell, and Peter Stone. 2018. 'Generative Adversarial Imitation from Observation'. *CoRR* abs/1807.06158. http://arxiv.org/abs/1807.06158.

Varela, Francisco J, Evan Thompson, and Eleanor Rosch. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. The MIT Press.

Vink, Roy, Maarten L Wijnants, Antonius HN Cillessen, and Anna MT Bosman. 2017. 'Cooperative Learning and Interpersonal Synchrony'. *Nonlinear Dynamics, Psychology, and Life Sciences* 21 (2): 189–215.

Walsh, Kevin S, David P McGovern, Andy Clark, and Redmond G O'Connell. 2020. 'Evaluating the Neurophysiological Evidence for Predictive Processing as a Model of Perception'. *Annals of the New York Academy of Sciences* 1464 (1): 242.

Wang, Jack M, Samuel R Hamner, Scott L Delp, and Vladlen Koltun. 2012. 'Optimizing Locomotion Controllers Using Biologically-Based Actuators and Objectives'. *ACM Transactions on Graphics (TOG)* 31 (4): 1–11.

Wang, Tingwu, Yunrong Guo, Maria Shugrina, and Sanja Fidler. 2020. 'UniCon: Universal Neural Controller For Physics-Based Character Motion'.

Welniarz, Quentin, Yulia Worbe, and Cecile Gallea. 2021. 'The Forward Model: A Unifying Theory for the Role of the Cerebellum in Motor Control and Sense of Agency'. *Frontiers in Systems Neuroscience* 15. https://doi.org/10.3389/fnsys.2021.644059.

Werling, Keenon, Dalton Omens, Jeongseok Lee, Ioannis Exarchos, and C Karen Liu. 2021. 'Fast and Feature-Complete Differentiable Physics Engine for Articulated Rigid Bodies with Contact Constraints'. In *Robotics: Science and Systems*. https://nimblephysics.org/.

Wolpert, Daniel M, Zoubin Ghahramani, and Michael I Jordan. 1995. 'An Internal Model for Sensorimotor Integration'. *Science* 269 (5232): 1880–82. https://doi.org/10.1126/science.7569931.

Won, Jungdam, Deepak Gopinath, and Jessica Hodgins. 2020. 'A Scalable Approach to Control Diverse Behaviors for Physically Simulated Characters'. *ACM Transactions on Graphics (TOG)* 39 (4): 33–1.

———. 2021. 'Control Strategies for Physically Simulated Characters Performing Two-Player Competitive Sports'. *ACM Trans. Graph.* 40 (4). https://doi.org/10.1145/3450626.3459761.

———. 2022. 'Physics-Based Character Controllers Using Conditional VAEs'. *ACM Transactions on Graphics (TOG)* 41 (4): 1–12. https://doi.org/10.1145/3528223.3530067.

Xie, Zhaoming, Sebastian Starke, Hung Yu Ling, and Michiel van de Panne. 2022. 'Learning Soccer Juggling Skills with Layer-Wise Mixture-of-Experts'. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Proceedings*, 1–9. https://doi.org/10.1145/3528233.3530735.

Xu, Pei, and Ioannis Karamouzas. 2021. 'A GAN-Like Approach for Physics-Based Imitation Learning and Interactive Character Control'. *Proceedings of the ACM on Computer Graphics and Interactive Techniques* 4 (3): 1–22. https://doi.org/10.1145/3480148.

Yamazaki, Tadashi, and William Lennon. 2019. ‘Revisiting a Theory of Cerebellar Cortex’. *Neuroscience Research* 148: 1–8. https://doi.org/10.1016/j.neures.2019.03.001.

Yan, Sijie, Yuanjun Xiong, and Dahua Lin. 2018. ‘Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition’. In *Thirty-Second AAAI Conference on Artificial Intelligence*. https://doi.org/10.1609/aaai.v32i1.12328.

Yin, Zhiqi, Zeshi Yang, Michiel Van De Panne, and KangKang Yin. 2021. ‘Discovering Diverse Athletic Jumping Strategies’. *ACM Transactions on Graphics (TOG)* 40 (4): 1–17. https://doi.org/10.1145/3450626.3459817.

Yin, Zhiqi, and KangKang Yin. 2020. ‘Linear Time Stable PD Controllers for Physics-Based Character Animation’. In *Computer Graphics Forum*, 39:191–200. Wiley Online Library.

Zhai, Chao, Francesco Alderisio, Piotr S\lowiński, Krasimira Tsaneva-Atanasova, and Mario di Bernardo. 2017. ‘Design and Validation of a Virtual Player for Studying Interpersonal Coordination in the Mirror Game’. *IEEE Transactions on Cybernetics* 48 (3): 1018–29.

Zhang, Haotian, Ye Yuan, Viktor Makoviychuk, Yunrong Guo, Sanja Fidler, Xue Bin Peng, and Kayvon Fatahalian. 2023. ‘Learning Physically Simulated Tennis Skills from Broadcast Videos’. *ACM Transactions Graphics* 42 (4): 14. https://doi.org/10.1145/3592408.

Zhang, He, Yuting Ye, Takaaki Shiratori, and Taku Komura. 2021. ‘ManipNet: Neural Manipulation Synthesis with a Hand-Object Spatial Representation’. *ACM Trans. Graph.* 40 (4). https://doi.org/10.1145/3450626.3459830.