Character Animation & Control

Mohamed Hassan, 10th September 2020

- Prototype vs product
- Animation expert?

Outlines

- What? Why? How?
    - What is character animation?
    - Why do we need this?
    - How can we do it?

What? Why? How?

What is character animation?

What is animation?

- Character animation is the task of making a character move in 2D or 3D space
- Physical vs simulated characters
- Control vs animation

What is animation?

- Character animation is the task of making a character move in 2D or 3D space
- Character interaction animation is when the movement involve another character(s) and/or object(s)

Classification

- Long term vs short term generation
- Controlled vs uncontrolled motion generation
- What type of control?
  - Compact user-input
  - mouse, keyboard, text, voice, IMU, tracking camera
- Physical or non-physical animation

What? Why? How?

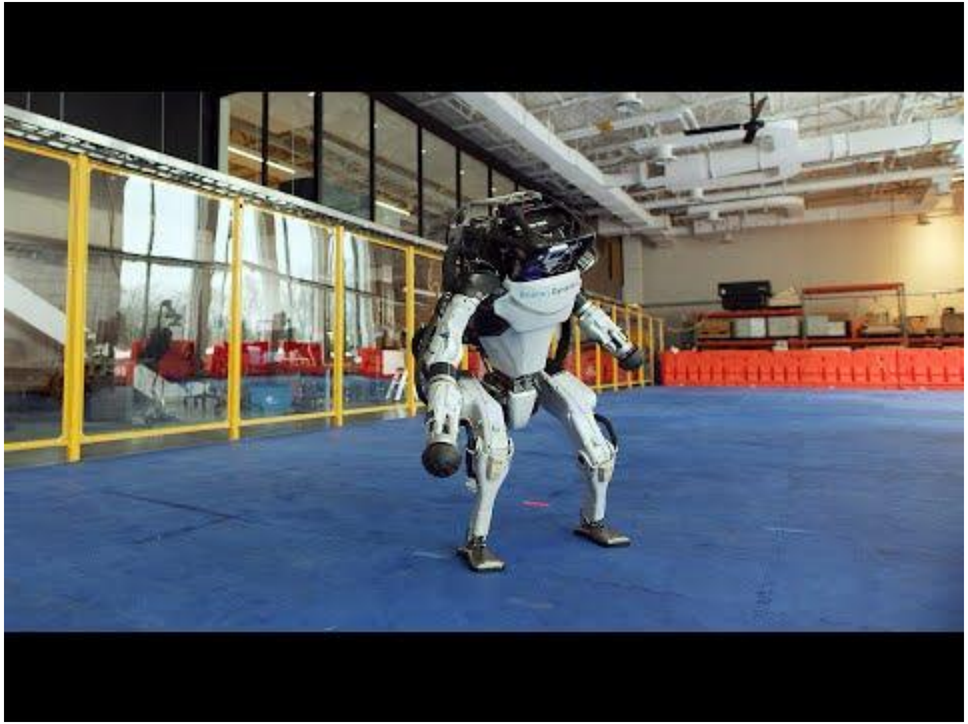
Why do we need animation?

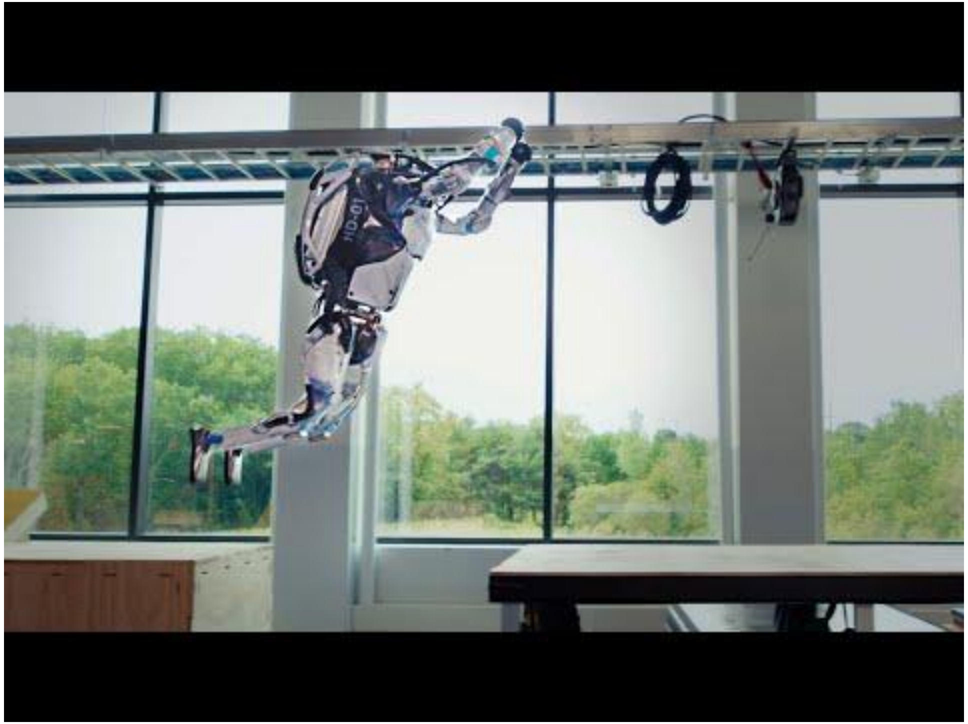
Why?

- Control a robot to achieve a goal
    - Carry a box
    - Open a door
- Future of communication
- Creating digital world/content aka Metaverse
    - Why? You decide on the goal
    - Communicating a better message
        - What is better than a 2d video?
        - Movies
    - Games
    - VR/AR
    - Synthetic data for training
    - Simulation and Optimization

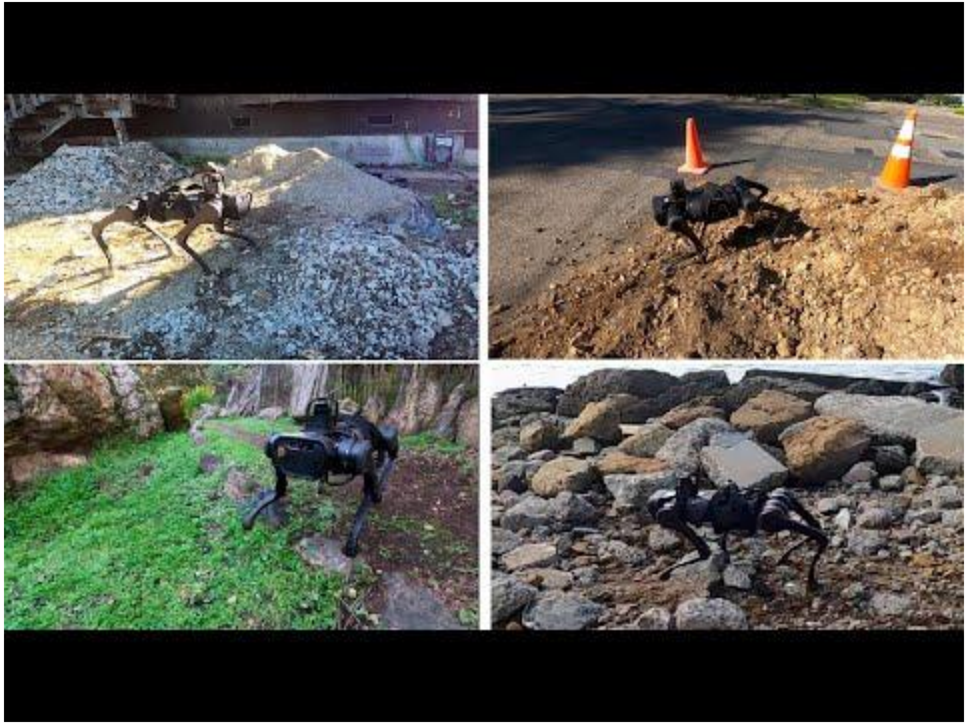
Examples

Robot Control





Credit to Boston Dynamics



Better Communication

- What is after video calls?
- Are you tired of looking at a screen?
- Can we bring people together in the same room?

Facebook Horizon Workrooms

- Bring your desk and laptop
- Sketch
- Handshake, high five?

Challenges

- How can we animate a virtual character?
- How can we create a virtual character? -> Another session
- Is the motion realistic?
- How to evaluate motion quality?
- Does it make you feel tired/exhausted?
- Would you use it everyday?

Microsoft Mesh

Metaverse

Collective virtual shared space, 3D internet

Optimizing workflow - BMW digital twin

Facebook horizon

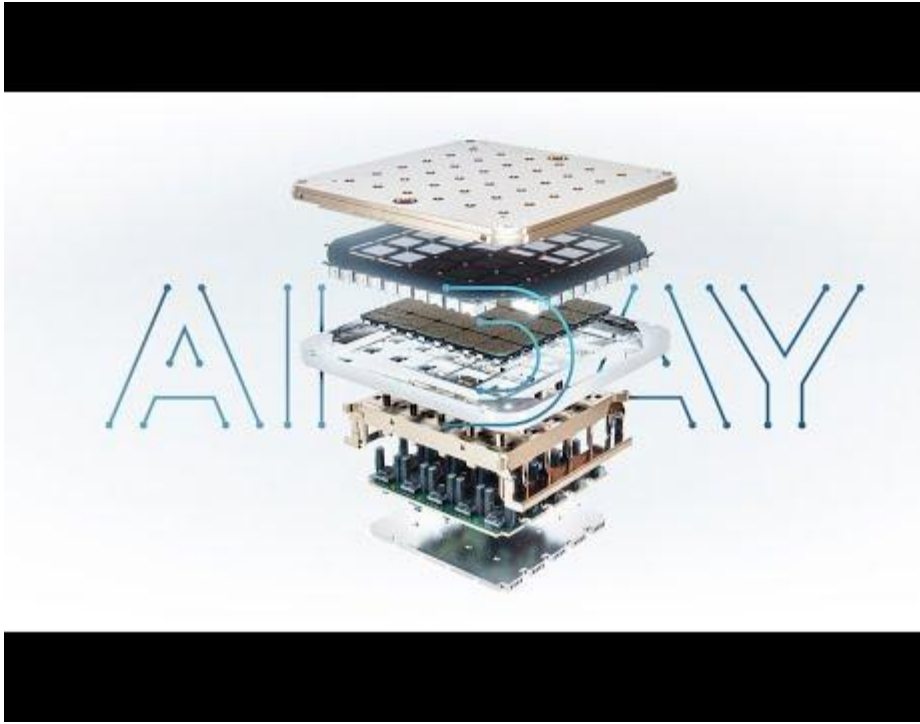
- Create your own world
- Invite others
- FB page vs FB world

Live Concert


CHARACTER CREATOR
NVIDIA OMNIVERSE
CREATING ANIMATED DIGITAL HUMANS FOR OMNIVERSE
USD

Simulation & Synthetic Data

Tesla Self-Driving Car



What? Why? How?

How can we do animation?

What do we want?

- Generated motion should be:
  - Realistic
  - Human-like
  - Diverse
    - How many types of motion can a human do?
- Algorithm\system should be
  - Real-time*
  - Easy to use and control

Is this hard? Why?

- How to represent/describe human motion?
- How to represent/describe rigid body motion?
  - Rigid body is a solid body with zero deformation
  - Position and orientation at each point of time
- Human motion representation
  - Skeletal tree
  - Root positions and orientation
  - Joint rotations
- Can we write equation for walking human?
  - Character state at each point of time

Is this hard? Why?

- Human motion is complex, diverse
- Our eyes are sensitive to artifacts
- Level of details:
  - Hand motion and finger articulation
  - Facial expressions

Pre-deep learning

- Re-visiting old ideas with new techniques
- History? Still used? Useful?
- Optimization-based

Methods

- Motion Capture (MoCap)
- Motion Editing and Retargetting
- Example-based methods
- Keyframe-based methods
- Autoregressive method
  - Physics-based
  - Data-based

- Motion Capture (MoCap)
- Motion Editing and Retargetting
- Example-based methods
- Keyframe-based methods
- Autoregressive method
  - Physics-based
  - Data-based

MoCap

- Motion capture is the process of recording the movement of objects or people.
- Technologies:
  - Optical markers: IR camera and reflective markers
  - Wearable sensors

MoCap

- Motion capture is the process of recording the movement of objects or people.
- Technologies:
    - Optical markers: IR camera and reflective markers
    - Wearable sensors

- Expensive
- Not scalable
    - Capture every actor
- Highest quality

- How to animate a new character(s)?
    - A dragon, a zombie, Gollum

Optimization Basics

- Find the largest possible rectangular area you can enclose, assuming you have X meters of fencing.
- *Objective function* A = W* L
- Constraint 2L + 2W = X
- Variables L, W
- Analytical solution
- Numerical solution
- Iterative methods
- Local minima
- Sensitive to initialization

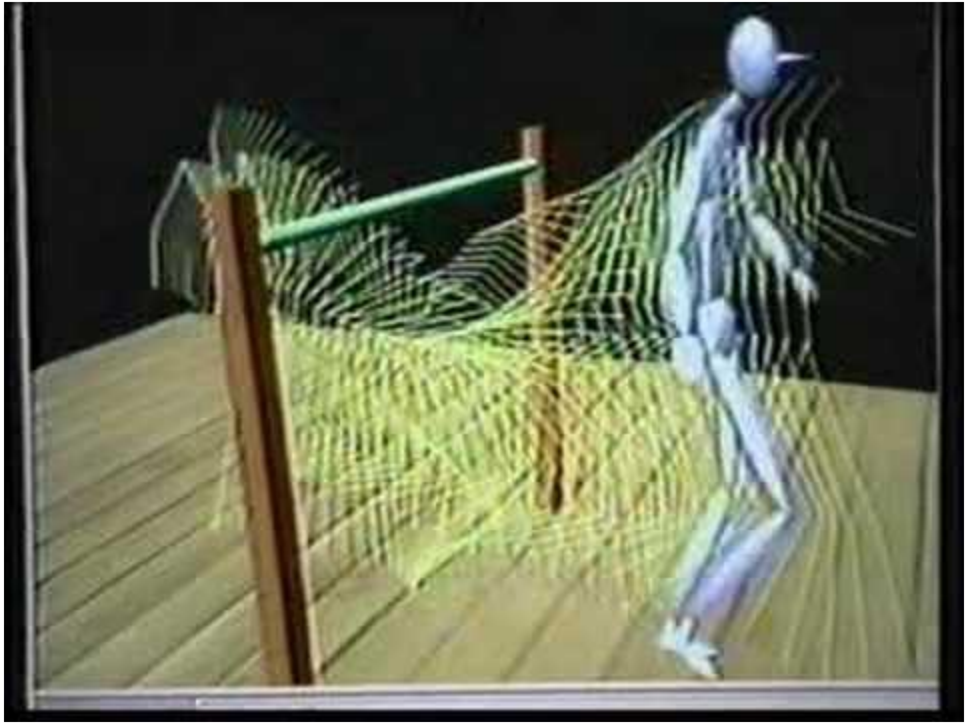
- Motion Capture (MoCap)
- Motion Editing and Retargetting
- Example-based methods
- Keyframe-based methods
- Autoregressive method
    - Physics-based
    - Data-based

Motion Editing

- Generate a new motion from a recorded one

Motion Editing

- Edit one frame
- Use IK solver
- IK is calculating joint parameters needed to place the end of a kinematic chain in a given position and orientation.
- Fit the new motion curve
- Requires manual work
- Limited to small edits
- Used by animation tools today
- Used today for data augmentation



LEE, J., AND SHIN, S. Y. 1999. A hierarchical approach to interactive motion editing for human-like figures

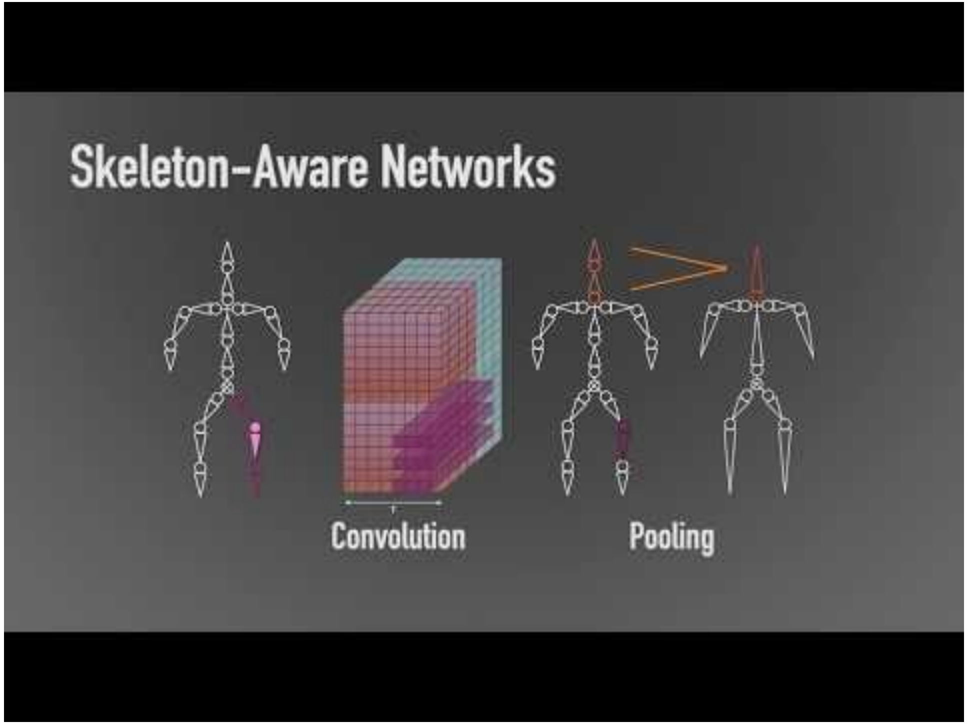
Motion Retargetting

- Adapting a MoCap from one character to another
- Different skeletal trees
- Different joint lenghts
- Specify some constraints which should be maintained (feet on the floor)
- Solve for the new motion considering the *whole* initial motion
- Solve: what is the best motion that meets a specified set of constraints?
- Limitation: Same structure, same number of bones

GLEICHER, M. 1998. Retargeting motion to new characters

Motion Retargetting

- Different skeleton
- NN to convert skeleton A to skeleton B
- Introduce skeleton pooling



Aberman et. al. 2020. Skeleton-Aware Networks for Deep Motion Retargeting

- Motion Editing and Retargetting
- Example-based methods
- Keyframe-based methods
- Autoregressive method
    - Physics-based
    - Data-based

Motion Graphs

- Given a corpus of MoCap
- Construct a directed graph “Motion Graph”
  - Original motion
  - Transitions
- Generate new motion by walks on the graph
- Identify “similar” clips and use blending to generate transitions
- How to build the graph?
  - Complex
- How to split motion clips?
- Where to insert edges/connections?
- How to search the graph?
- How to generate transitions?
- High quality

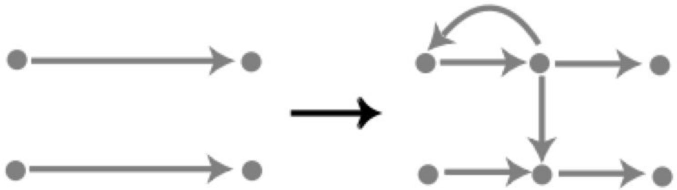
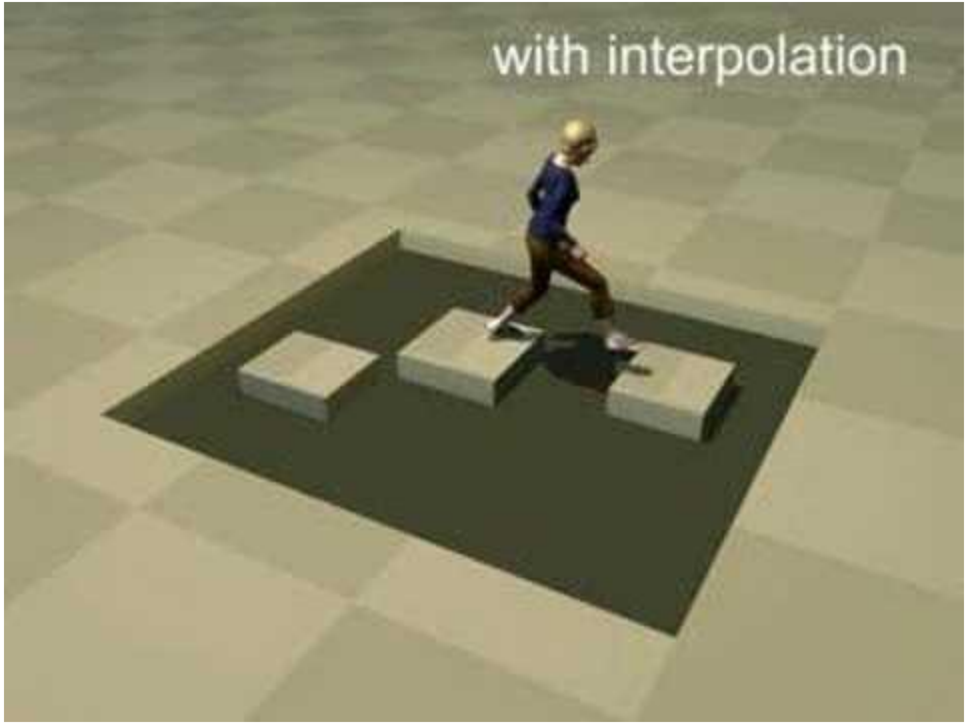


Figure 2: Consider a motion graph built from two initial clips. (top) We can trivially insert a node to divide an initial clip into two smaller clips. (bottom) We can also insert a transition joining either two different initial clips or different parts of the same initial clip.

KOVAR et. al. 2008. Motion graphs

Motion Graphs

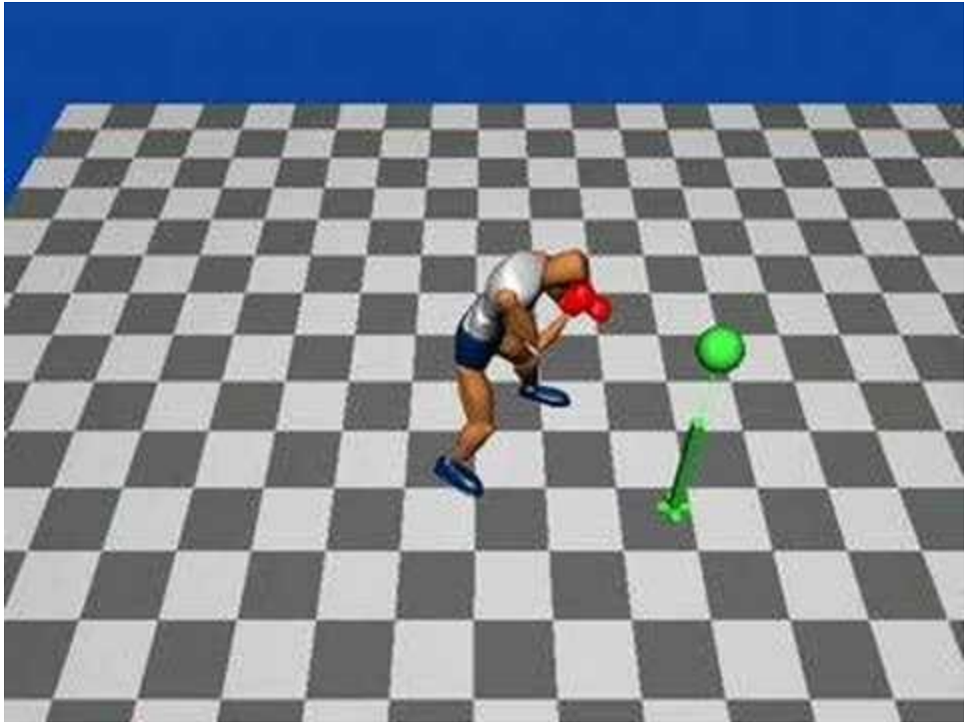
- Find motion which satisfy some constraints
- Search within a corpus of similar data
- Work for *similar* environments



Alla Safonova and Jessica K. Hodgins. 2007. Construction and Optimal Search of Interpolated Motion Graphs.

Motion Graphs

- Find motion which satisfy some constraints
- Given current state(pose, root) and the goal state, what is the most similar motion clip?



Jehee Lee and Kang Hoon Lee. 2004. Precomputing Avatar Behavior from Human Motion Data

Pushing People Around

- Given current state and input force, what is the most similar motion clip?

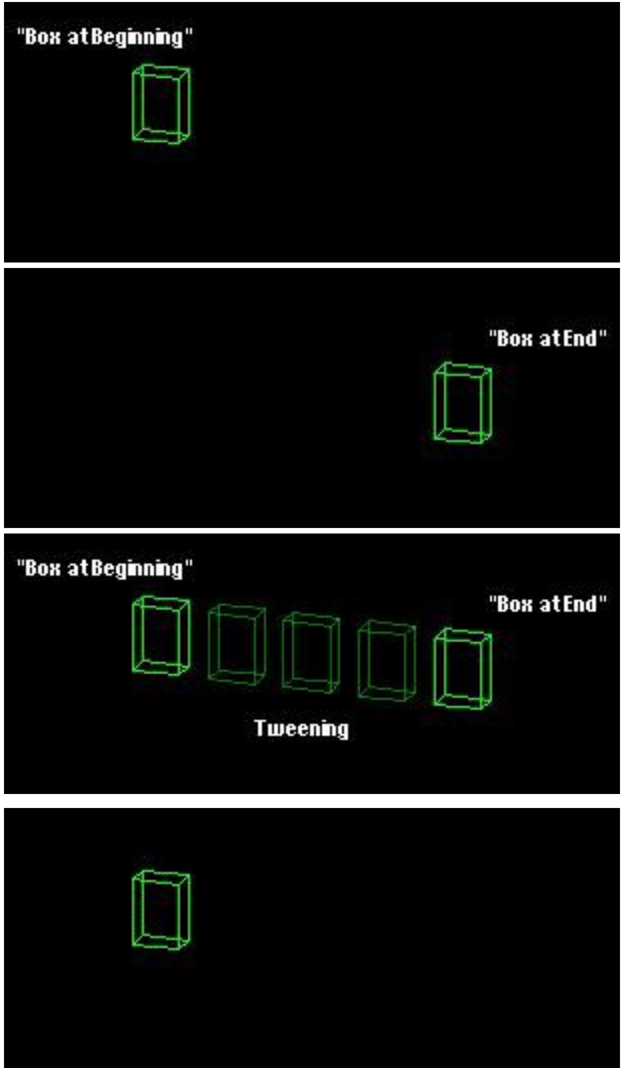
Arikan, O., Forsyth, D. A., O'Brien, J. F. 2005. Pushing People Around

Learned Motion Matching

Holden et. al.. 2020. Learned Motion Matching

- Motion Editing and Retargetting
- Example-based methods
- Keyframe-based methods
- Autoregressive methods
  - Physics-based
  - Data-based

Key-frame based method



https://web.arch.virginia.edu/arch545/handouts/keyframing.html

Motion In-betweening

- Manually define keyframes
- Generate the in-between motion
- Multimodal
- Divergence
- From Ubisoft, creator of Assassin's Creed

Harvey et. al. 2020. Robust Motion In-betweening

- Motion Editing and Retargetting
- Example-based methods
- Keyframe-based methods
- Autoregressive methods
  - Physics-based
  - Data-based

Autoregressive method

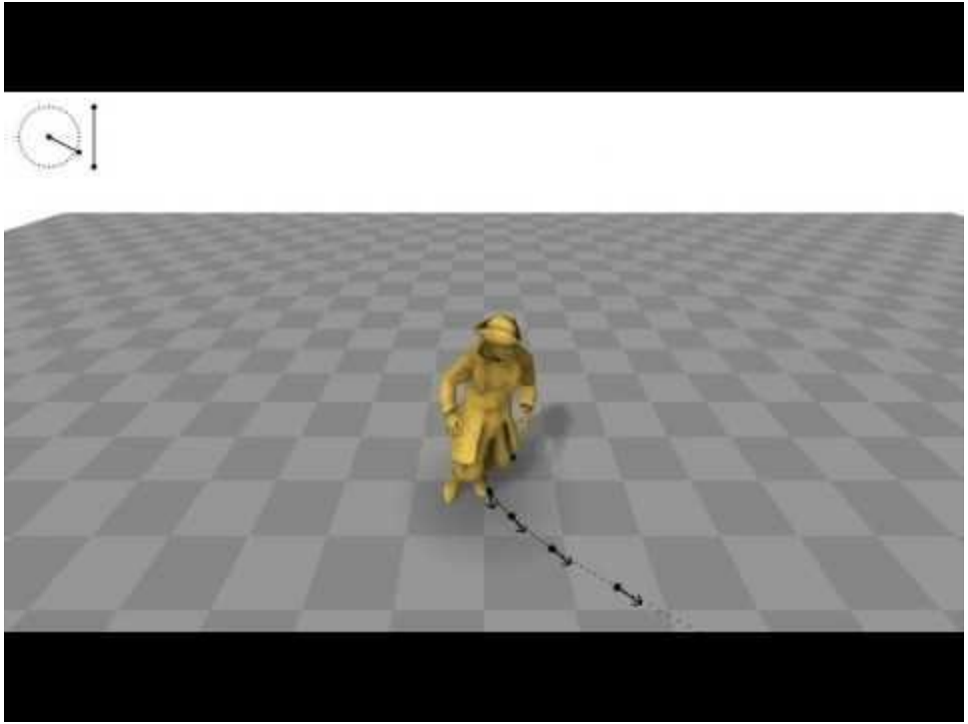
- Not replaying/editing existing example
- Generate new motion
- Rich and diverse motion
- Two types:
    - Data-based
    - Physics-based

- Motion Editing and Retargetting
- Example-based methods
- Keyframe-based methods
- Autoregressive methods
  - Data-based
  - Physics-based

- Vanilla Models
- Phase-Functioned Neural Network (PFNN)
- Mixture of experts
- Scheduled sampling

Data-based Methods

- Classic autoregressive models(RNN, LSTM,...)
  - Smooth results
  - Fail for long motion sequences
  - Die out or explode



Phase-Functioned Neural Network (PFNN)

- Intuition: differentiate between modes
- Wights of NN are different for each motion mode
- Different NN for each motion cluster
- More representation power
- phase==cycle==progression
- Compute the NN weights based on the phase

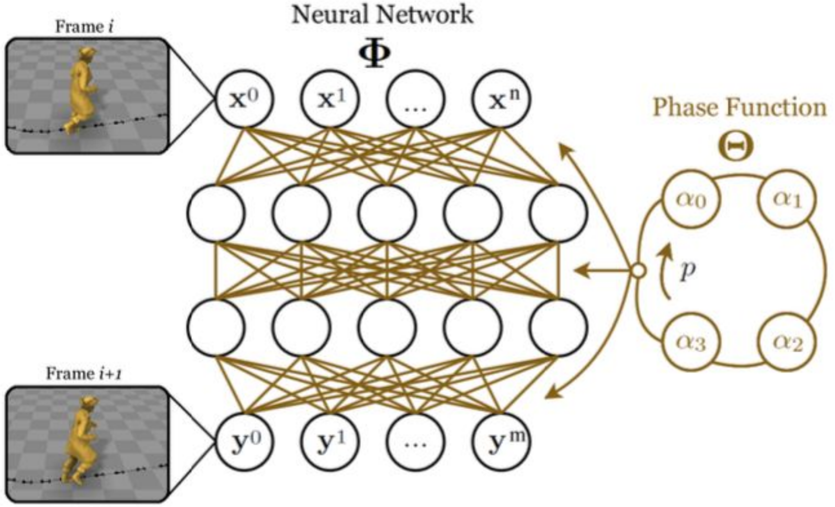
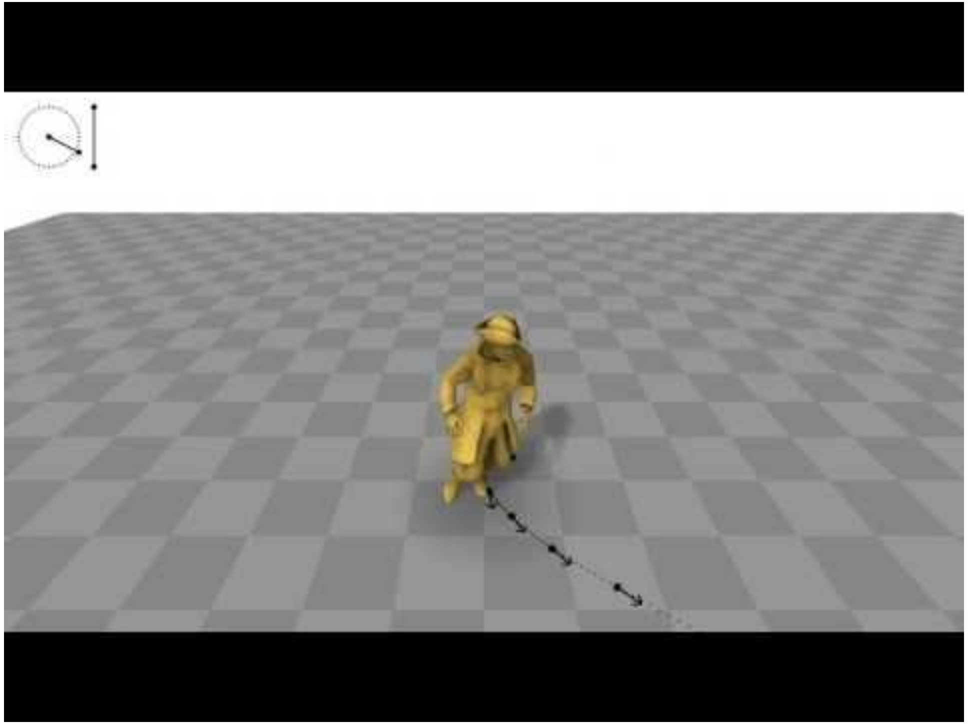


Fig. 2. Visual diagram of Phase Functioned Neural Network. Shown in yellow is the cyclic *Phase Function* - the function which generates the weights of the regression network which performs the control task.

Holden et. al. 2017. Phase-Functioned Neural Networks for Character Control

Phase-Functioned Neural Network (PFNN)

- 1M views on 09 Sep 2021



Holden et. al. 2017. Phase-Functioned Neural Networks for Character Control

Mode-Adaptive Neural Networks for Quadruped Motion Control (MANN)

- No manual phase labeling
- Mixture of experts
- *Predict* blending weights

Zhang et. al. 2018. Mode-Adaptive Neural Networks for Quadruped Motion Control

Mode-Adaptive Neural Networks for Quadruped Motion Control (MANN)

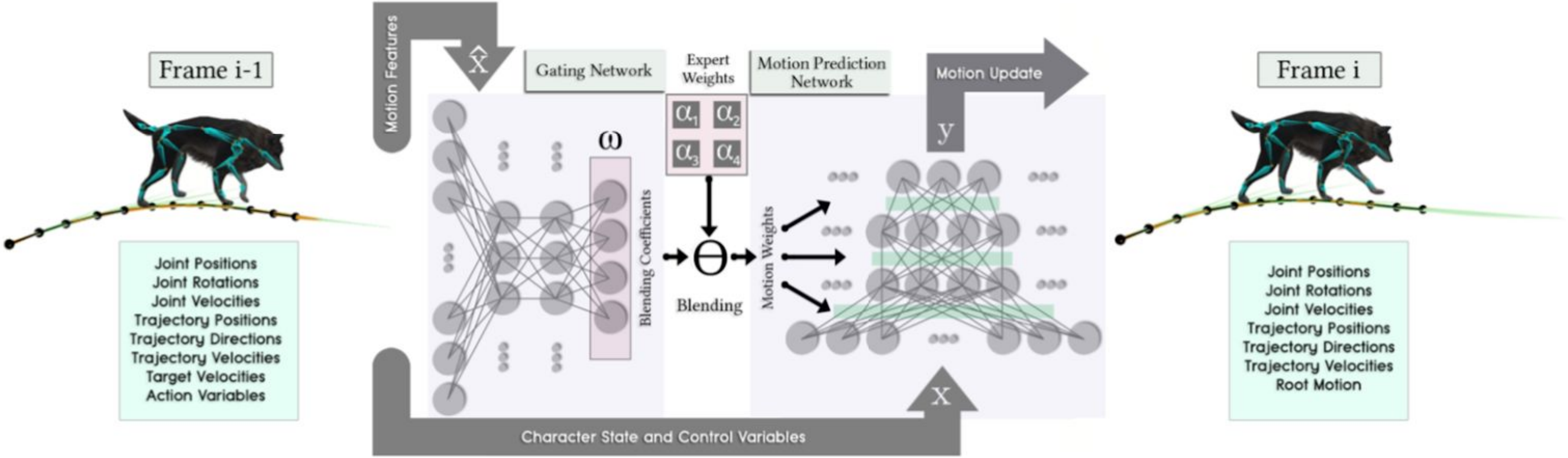


Fig. 3. The architecture of our neural network composed of the gating network and the motion prediction network. The gating network takes as input the current end effector velocities of the feet, the desired velocity and the action vector. The motion prediction network takes as input the posture and trajectory control variables from the previous frame, and predicts the updated posture and trajectory for the current frame.

Zhang et. al. 2018. Mode-Adaptive Neural Networks for Quadruped Motion Control

Mode-Adaptive Neural Networks for Quadruped Motion Control (MANN)

Zhang et. al. 2018. Mode-Adaptive Neural Networks for Quadruped Motion Control

Neural State Machine for Character-Scene Interactions (NSM)

- Focus on interaction
- 5 different actions with one NN
  - Idle, walk, run, carry and sit
- PFNN + MANN
- Predict phase + Mixture of experts

Starke et. al. 2019. Neural State Machine for Character-Scene Interactions (NSM)

NSM

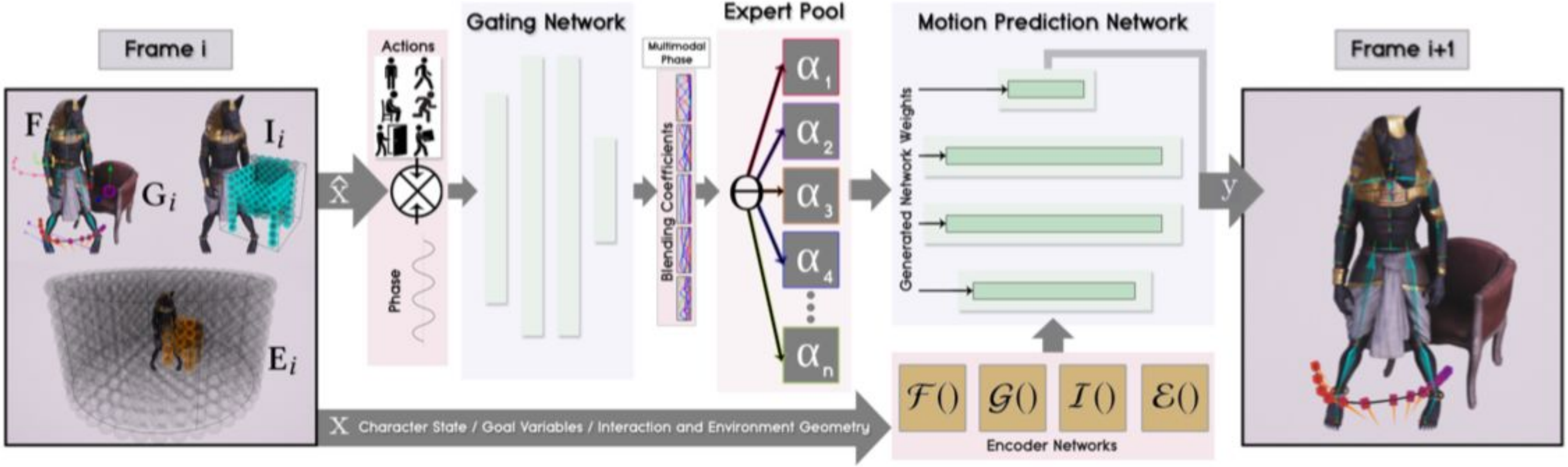


Fig. 2. The architecture of our system composed of the gating network and the motion prediction network. The gating network takes as input a subset of parameters of the current state and the goal action vector to output the expert blending coefficients which is then used to generate the motion prediction network. The motion prediction network takes as input the posture, trajectory control variables and the goal parameters from the previous frame, and predicts those variables for the current frame.

Starke et. al. 2019. Neural State Machine for Character-Scene Interactions (NSM)

NSM

Starke et. al. 2019. Neural State Machine for Character-Scene Interactions (NSM)

Scheduled Sampling

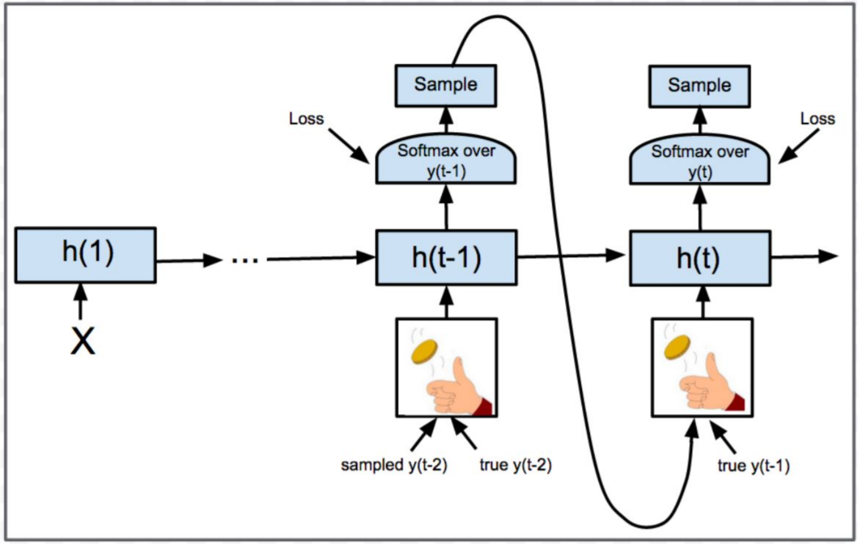


Figure 1: Illustration of the Scheduled Sampling approach, where one flips a coin at every time step to decide to use the true previous token or one sampled from the model itself.

Bengio et. al. 2015. Scheduled Sampling for Sequence Prediction with Recurrent Neural Networks

SAMP

Hassan et. al. 2021. Stochastic Scene-Aware Motion Prediction

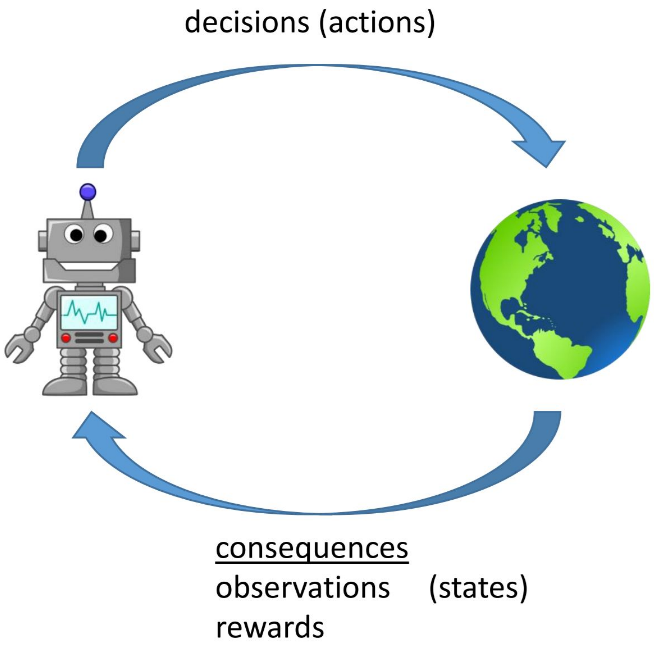
- Motion Editing and Retargetting
- Example-based methods
- Keyframe-based methods
- Autoregressive methods
    - Data-based
    - Physics-based

Physics-based Animation

- Simulating physically plausible behaviour
- Simulated physical environment
- Approximations
- RL based

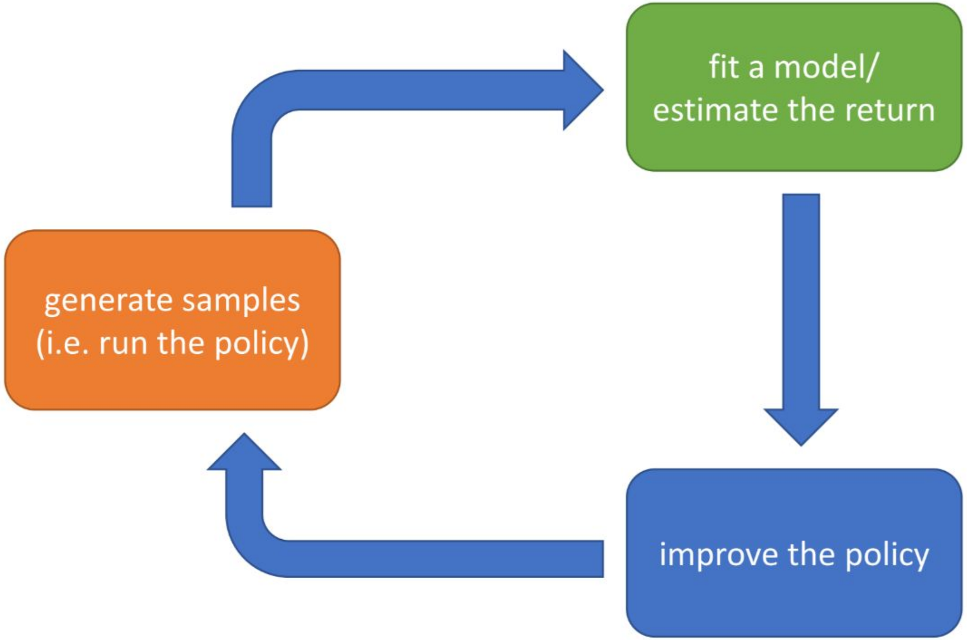
RL Basics

- Approach for learning decision making and control from experience
- Ground truth answer is not known, only know if we succeeded or failed (reward)



Credit to Sergey Levine

The anatomy of a reinforcement learning algorithm



Credit to Sergey Levine

Animating a Humanoid Character/Agent

- Humanoid Agent
- Physical environment
  - NVIDIA PhysX, Bullet
- Observation/State
  - Root position and orientation
  - Joint rotations
- Reward
  - Ambiguous
  - Moving forward
  - Not falling
- Train a policy
  - Training algorithm

Policy Gradient Algorithm

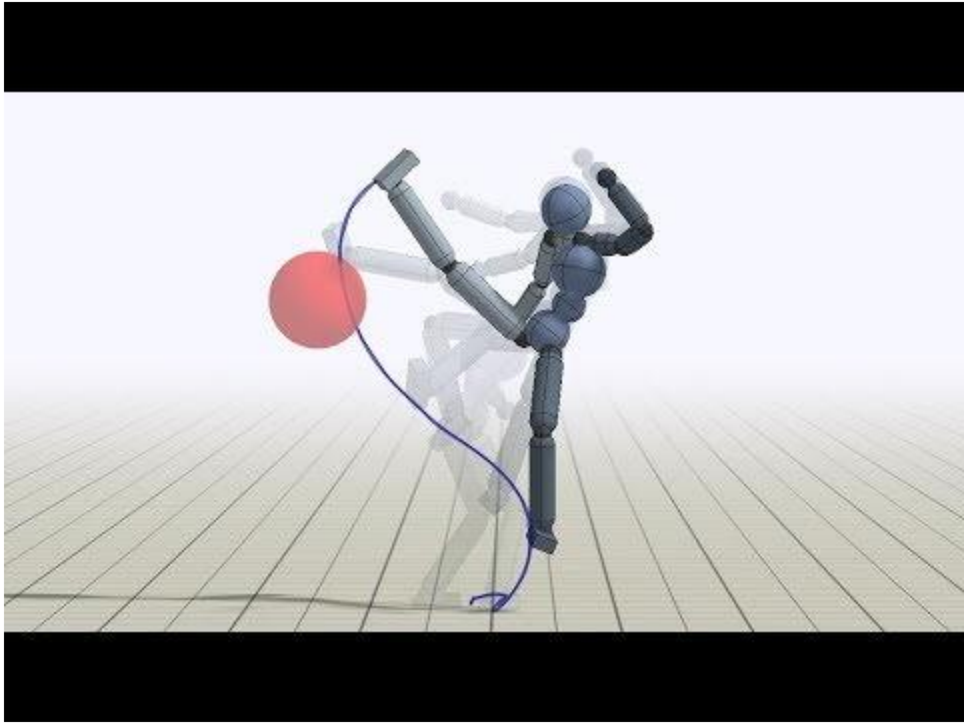
1. Initialize the policy parameter θ at random.
2. Generate one trajectory on policy π_θ: $S_1, A_1, R_2, S_2, A_2, \ldots, S_T$.
3. For t=1, 2, ... , T:
    1. Estimate the the return G_t;
    2. Update policy parameters: $\theta \leftarrow \theta + \alpha\gamma^t G_t \nabla_\theta \ln \pi_\theta(A_t|S_t)$

Lilian Weng blog

- How to define the reward?
- NN find peculiar ways to cheat
- Unnatural motion
    a. Reference motion (Deep Mimic)
- Little training data needed

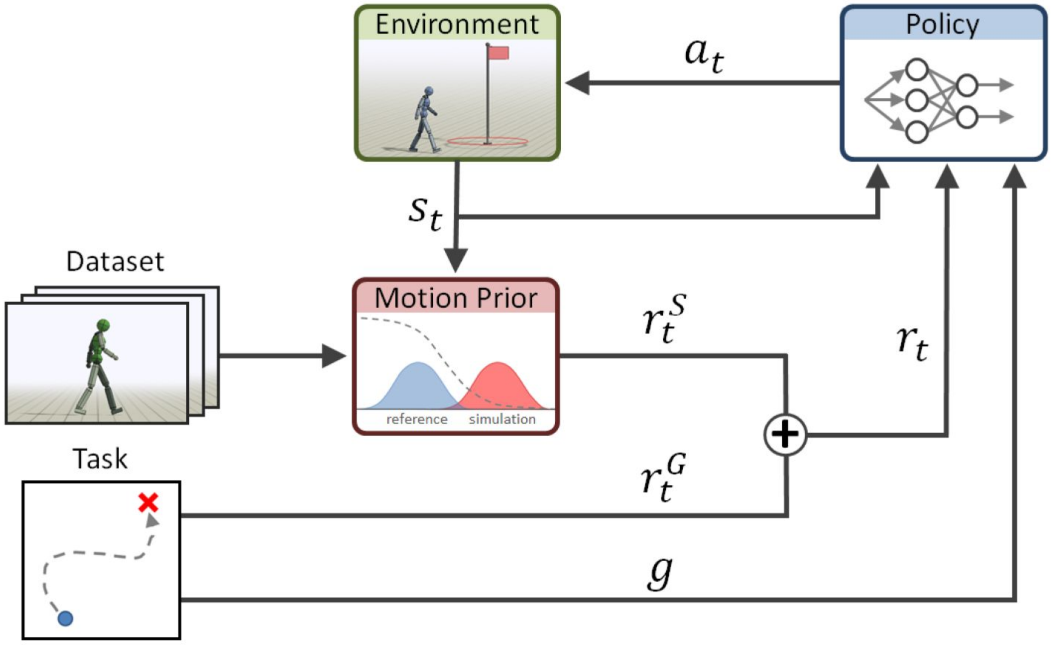
Deep Mimic

- Achieve a task while mimicking the reference motion
- Policy for each reference motion
- How to compose different skills/motions?



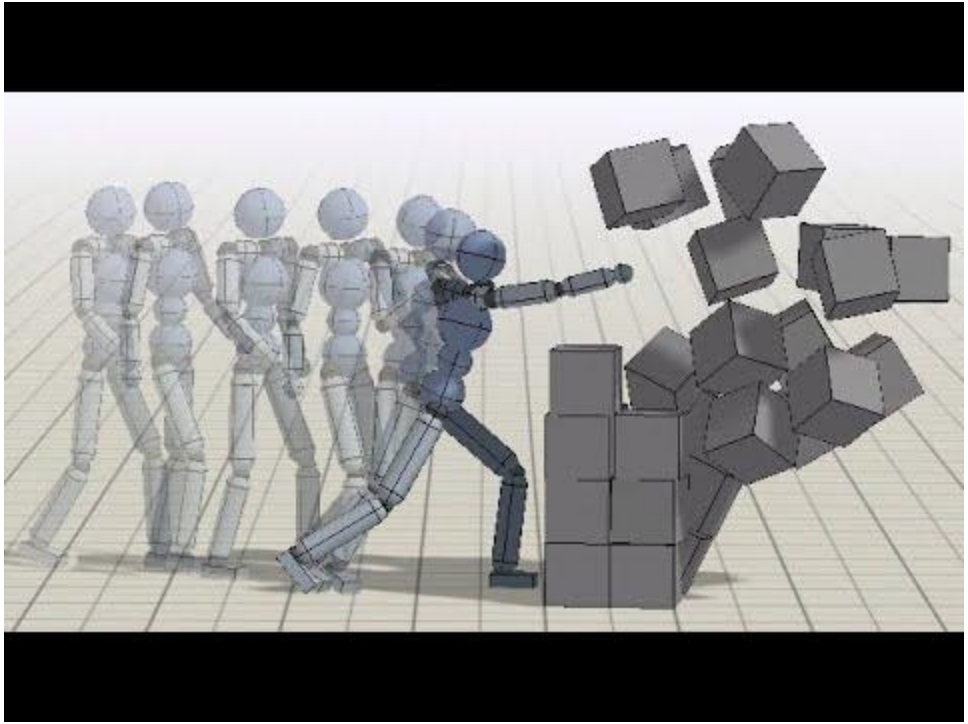
Peng et. al. 2018. DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills

AMP: Adversarial Motion Priors

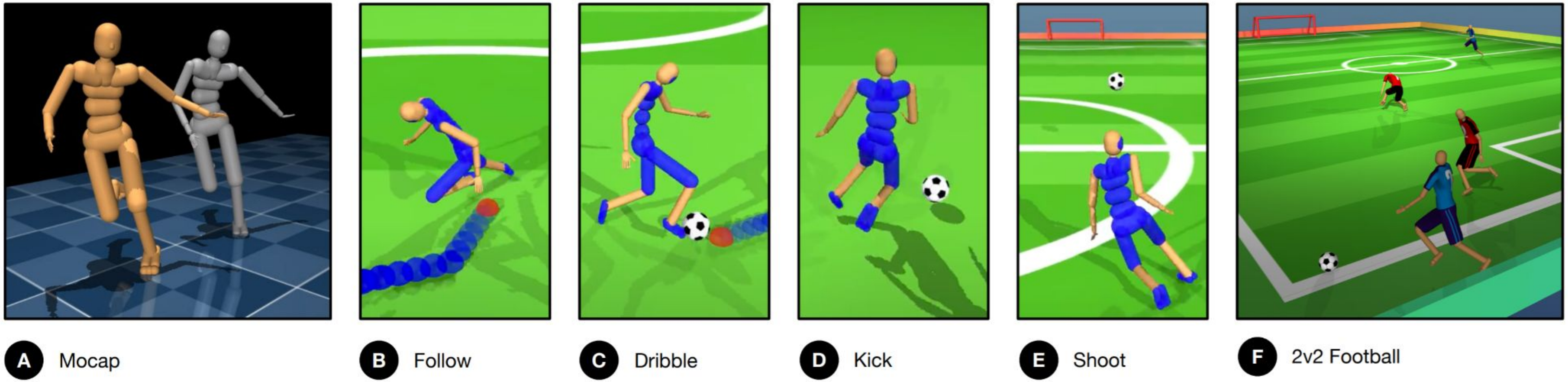


Peng et. al. 2021. AMP: Adversarial Motion Priors for Stylized Physics-Based Character Control

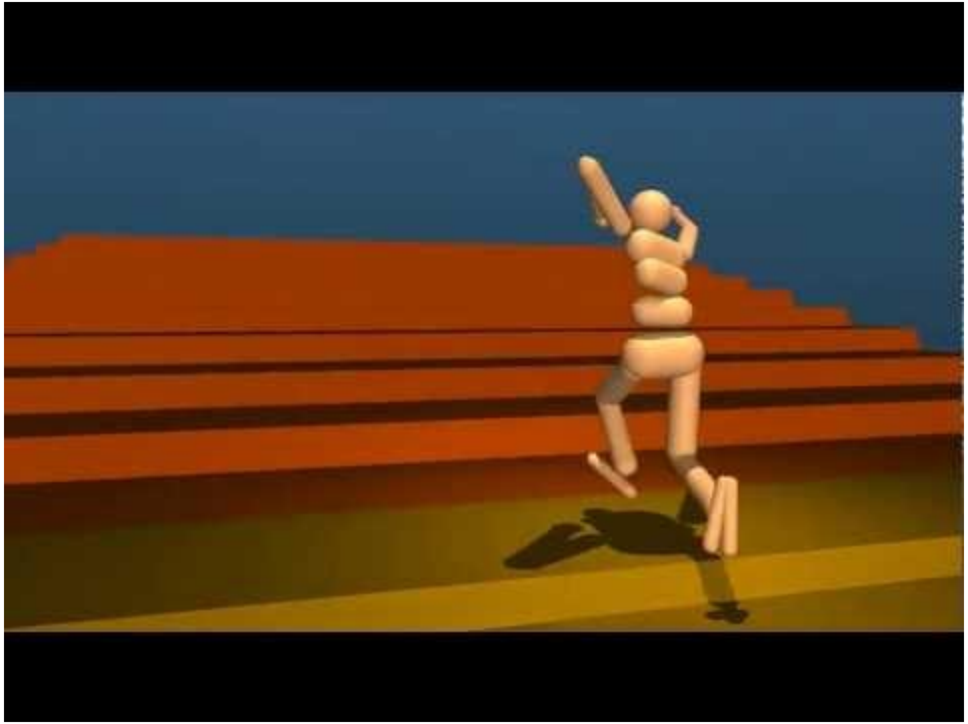
- Single policy for multiple skills



From Motor Control to Team Play

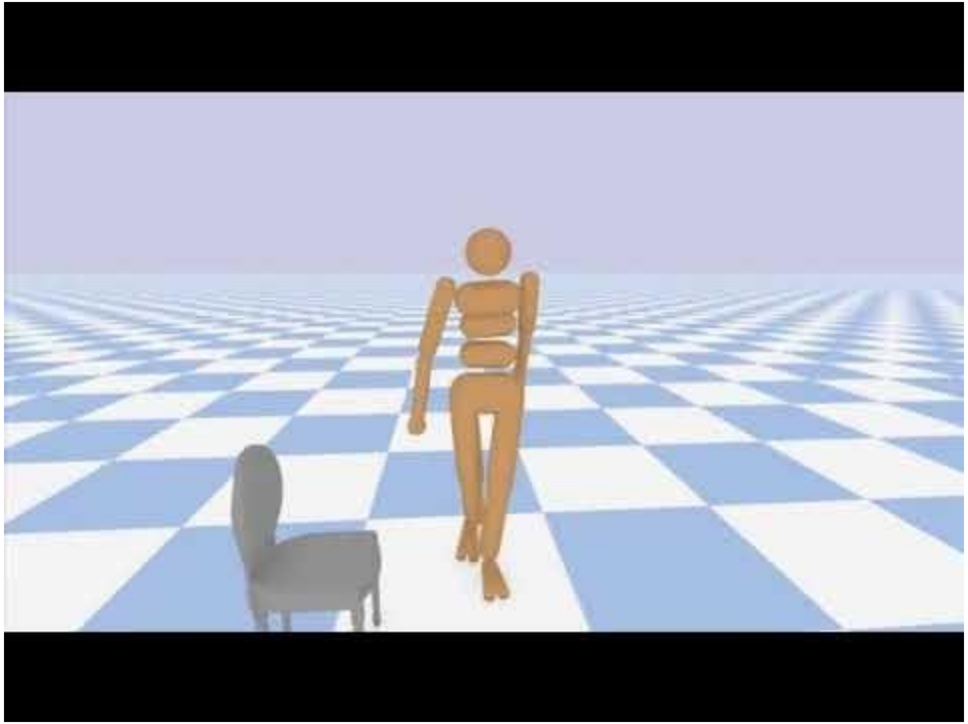


From Motor Control to Team Play in Simulated Humanoid Football.



Hesse et. al. 2017. Emergence of Locomotion Behaviours in Rich Environment

Learning to Sit

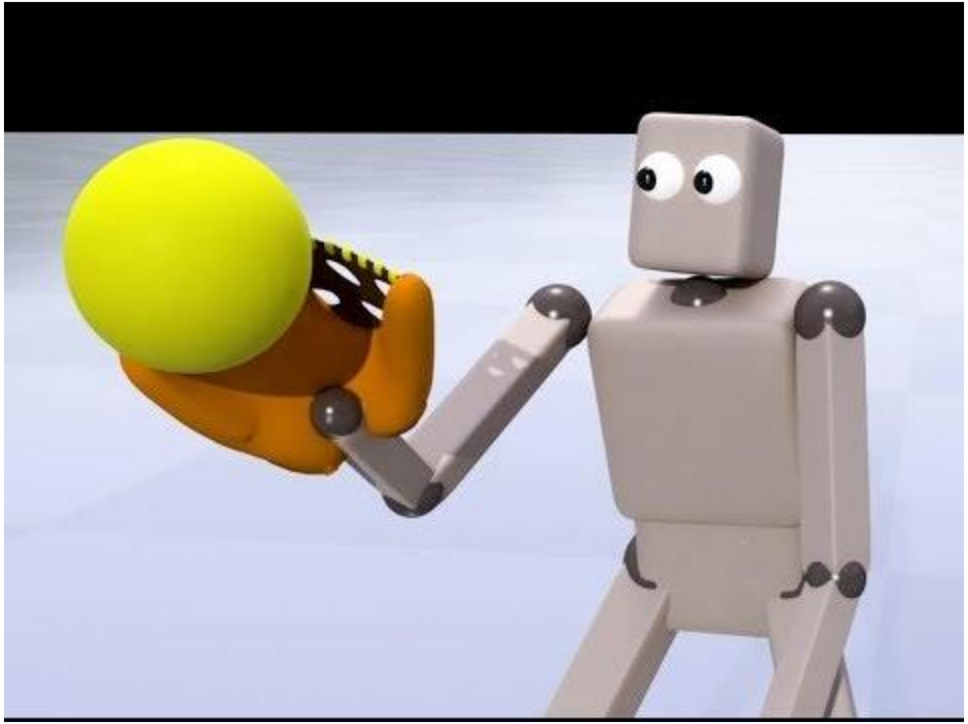


Chao et. al. 2021. Learning to Sit: Synthesizing Human-Chair Interactions via Hierarchical Control

Catch & Carry

- How would a robot perceive the environment?
- Transferability to real robots
- Can a robot get a 3D mesh of the environment?
- First-person perception, touch sensors and egocentric vision

Merel et. al. 2020. Catch & Carry: Reusable Neural Controllers for Vision-guided Whole-body Tasks



Eom et. al. 2020. Model Predictive Control with a Visuomotor System for Physics-based Character Animation

Thanks!