Improving Clinical Care of Pediatric Cerebral Palsy Patients with Inverse Reinforcement Learning

Mariah Schrum and Matthew Gombolay

Georgia Institute of Technology, Atlanta, Georgia

{mschrum3@,matthew.gombolay@cc.}gatech.edu

Keywords— **cerebral palsy, gait analysis, inverse reinforcement learning**

Cerebral palsy (CP) patients exhibit pathological gait patterns as a result of a variety of neuromuscular defects. These gait patterns are typically used to inform therapeutic treatment, yet outcomes vary significantly among individuals within a gait class. We investigate inverse reinforcement learning as an approach to discover latent features of CP gait to help clinicians better understand an individual patient's pathology and aid in clinical decision making. Furthermore, we develop deep reinforcement learning techniques that can prescribe ways in which a patient's gait might be altered to help a patient better achieve their ideal gait.

A. *Introduction*

Cerebral palsy occurs due to brain injury during early development and often results in impaired motor skills [1]. This condition manifests itself in many forms. Symptoms include muscle spasticity, lack of postural control, and other biomechanical deficiencies. Physical therapy is commonly prescribed to improve a patient's strength and mobility. Because biomechanical deficiencies and movement disorders are common across this patient population, characteristic gait abnormalities are often used to inform therapy. For example, *crouch gait*, which is characterized by excessive dorsiflexion at the ankle, may be treated with a ground reaction ankle foot orthosis or strengthening exercises [2]. If these treatments are not successful, therapists iteratively apply alternative therapies to home in on the best approach for an individual patient.

The variability in clinical results makes treating this pathology particularly challenging for therapists. Because two individuals exhibiting similar gait patterns (e.g., *crouch gait*, *equinus*, etc.) do not necessarily benefit from similar treatments, knowledge about a patient's gait class is not always enough information for therapists to determine the most effective treatment. Without tools to gain further understanding of an individual's specific pathology, trial and error is often a therapist's best method for determining a treatment plan. It is hypothesized that this variability is resultant from the fact that CP patients attempt to achieve different optimality criteria, subject to their own biomechanical constraints [4,5]. However, these goals are typically latent and difficult to reveal.

Previous research has indicated that human gait patterns are selected based on certain criteria that the body is attempting to optimize. Alexander [4] proposes that humans optimize for energy, while others have proposed minimization of fatigue and maximization of symmetry. Because CP patients differ biomechanically from healthy adults, comparing CP gait patterns to typical gait alone is not an adequate indicator of improvement and instead a better approach may be to craft therapies to aid the patients in achieving these optimality criteria [5]. Knowledge about the optimality principles with which CP patients are complying may serve as a better indicator for therapists when selecting a treatment plan. For instance, if the patient's specific gait pattern is a result of the body attempting to minimize asymmetry, the therapist may provide techniques that help the patient to accomplish this goal.

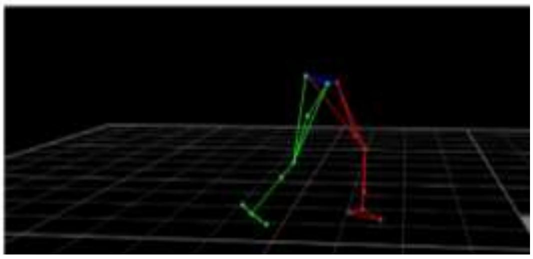

Fig 1. Motion capture data from cerebral palsy subject

Jeng et al. [5] investigated which criteria CP children optimize by having children with CP walk at their preferred frequency (strides per minute) and measuring how certain criteria changed when the children were forced to walk at faster or slower frequencies. These results were compared to the behaviour of typical children. The findings indicate that typical children attempt to minimize physiological costs, asymmetry in lower limb movement, and variability of inter and intralimb coordination. It was determined that CP patients may optimize for all or only some of these criteria. Thus, it is hypothesized that CP children may be employing specific optimality strategies to adapt to their varying biomechanical limitations. An understanding of these strategies may help guide therapists in determining the most effective treatment plan to help patients achieve their ambulatory goals.

B. *Inverse Reinforcement Learning*

Exploring the optimality principles of CP patients shows clinical promise. Thus, our goal is to determine the optimality criteria of individual CP patients to provide therapists with a more informative profile so that they can construct a treatment plan that has a higher probability of success. We build on the previous work of Jeng et al. [5] and propose using mathematical techniques to determine these optimality criteria.

We frame the problem of determining the optimality criteria in terms of inverse reinforcement learning (IRL). IRL is the problem of finding a reward function that explains a given set of trajectories. The basis function for the reward is typically assumed to be a weighted linear combination of the features associated with the trajectories. The objective of IRL is to recover the weights associated with the features of the trajectories and construct a policy to maximize this reward function. Framing the problem of IRL as a solution to a Markov Decision Process provides a mathematical model of the decision-making process and provides constraints on feasibility and practicality of the actor's decisions. However, recovering the exact reward function is not a fully defined problem since the true basis function is unknown and there are many optimal policies which can explain given trajectories. To attempt to overcome some of these challenges, Zietbart et al [6] propose a maximum entropy formulation in which trajectories (i.e., sequences of state action pairs) are assumed to be distributed according to Eq 1. This maximum entropy formulation resolves this ambiguity in choosing distributions over behaviours and allows for uncertainty when the trajectories are imperfect. Here τ represents the trajectories, θ the weights to be recovered, T the transition distribution, *s* the state or joint configuration, and *a* the action or muscle excitations. The function Z is the partition function or normalization term.

$$P(\tau|\theta,T) = \frac{e^{\theta f_\tau}}{Z(\theta,T)} \prod_{s_{t+1},s_t,a_t \, \epsilon \, \tau} P_T(s_{t+1}|a_t,s_t) \qquad \text{(Eq 1)}$$

Recovering the weights of the features can be achieved by maximizing the likelihood of the trajectories subject to the principle of maximum entropy. Several related approaches have also been proposed which accommodate for continuous time and action spaces [7], [8]. For example, Aghasadeghi and Bretl [9] suggest using a path integral formulation and solving a similar maximum likelihood problem. Because the gait trajectories in our case consist of human gait data, the trajectories will be noisy. Thus, a maximum entropy IRL formulation will be a good candidate for inferring the reward function that explains the CP gait data.

To determine patient-specific reward functions, our clinical partners at Georgia State University have collected lower body motion capture data and EMG data from children with CP. A sample of the motion capture data is picture in Fig 1. Given this set of motion capture and EMG data of CP patients, we propose to utilize IRL techniques to determine the criteria CP patients are optimizing. This kinematic data consisting of joint states and muscle activations can be thought of as demonstrated trajectories in terms of IRL. Our objective is to learn the underlying reward function that explains these demonstrated trajectories. This reward function will elucidate the optimality principles which are governing the kinematics of a patient in an interpretable way and provide important insight for therapists when determining a treatment plan.

C. Machine assisted therapy

Having inferred the goals of an individual CP patient, the next step is to determine how to help the patient better achieve these goals. Therapists would benefit from assistance in understanding the minimal alterations that could be made to a patient's gait to achieve maximum ambulatory benefit. To explore mathematically ways in which a patient's gait may be improved, we will employ reinforcement learning techniques.

Given the biomechanical constraints and other limitation of an individual CP patient, we want to determine the optimal policy with respect to the patient's goals for the patient to follow when ambulating. To do so, we leverage the Stanford OpenSim dynamic musculoskeletal physics simulator to infer patients' goals given their kinematic limitations. Currently, we are exploring techniques with the 2D simulation and will build to the 3D simulation. With this framework, we can apply biomechanical constraints to the simulator that match those of a CP patient. To date, we have trained the simulator to walk using Deep Deterministic Policy Gradient [10], as pictured in Fig 2. We do so for both a typical individual and an individual with a biomechanical impairment common in CP patients, i.e., weak calf muscles.

Preliminary results show we can learn an ambulatory policy on both the typical simulation and simulation with weak calf muscles. Further training and improvements need to be made to learn the most efficient gait. However, our preliminary results suggest that this is a promising endeavour. A video of

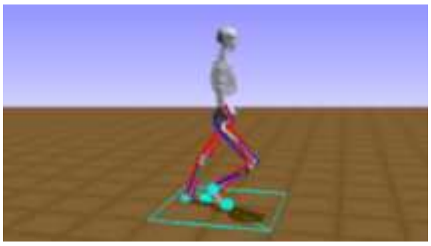

Fig 2. DDPG applied to Stanford OpenSim Model

preliminary results can be found here: http://tiny.cc/1xrv4y. In practice, these results will provide clinicians with suggestions about how a patient's gait may be augmented to better enable the patient to achieve his or her goals.

D. Future Work

Our clinical collaborators are continuing to collect gait data from CP patients. As we gain further data, we will begin applying IRL and reinforcement learning techniques to infer patients' latent goals and what therapies (gait augmentations) could be affected to most improve ambulatory benefit. The aim is to aid human therapists in clinical decision making by tailoring therapies for each patient to achieve the best outcomes.

E. ACKNOWLEDGEMENTS

This work is supported by the NSF Accessibility, Rehabilitation and Movement Science Fellowship under Grant # 1545287

References

[1] Rosenbaum P, Paneth N, Leviton A, et al. A report: the definition and classification of cerebral palsy April 2006. *Developmental medicine and child neurology Supplement.* 2007;109:8-14..

[2] "Classification of Gait Patterns in Cerebral Palsy." *Physiopedia*,www.physiopedia.com/Classification_of_Gait_Patterns_in_Cerebral_Palsy.

[3] Zhang, Yanxin, and Ye Ma. "Application of Supervised Machine Learning Algorithms in the Classification of Sagittal Gait Patterns of Cerebral Palsy Children with Spastic Diplegia." *Computers in Biology and Medicine*, vol. 106, 2019, pp. 33–39., doi:10.1016/j.compbiomed.2019.01.009.

[4] Alexander, R. Mcn. "The Gaits of Bipedal and Quadrupedal Animals." *The International Journal of Robotics Research*, vol. 3, no. 2, 1984, pp. 49–59., doi:10.1177/027836498400300205.

[5] Jeng, Sun-Fang, et al. "Self-Optimization of Walking in Nondisabled Children and Children with Spastic Hemiplegic Cerebral Palsy." *Journal of Motor Behavior*, vol. 28, no. 1, 1996, pp. 15–27., doi:10.1080/00222895.1996.9941729.

[6] Ziebart, B., Maas, A., Bagnell, J. A., and Dey, A. K. Maximum entropy inverse reinforcement learning. In AAAI Conference on Artificial Intelligence, 2008.

[7] Levine, S. and Koltun, V. Continuous inverse optimal control with locally optimal examples. In International Con- ference on Machine Learning (ICML), 2012.

[8] Chelsea Finn, Sergey Levine, and Pieter Abbeel. 2016. Guided cost learning: deep inverse optimal control via policy optimization. In Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48 (ICML'16).

[9] Aghasadeghi, N. and Bretl, T. Maximum entropy inverse reinforcement learning in continuous state spaces with path integrals. In International Conference on Intelligent Robots and Systems (IROS), 2011

[10] Ljungstrom, Mattias. "An asynchronous DDPG Reinforcement with multiple Actor-Critic pairs". (2018). Github repository: https://github.com/mattiasljungstrom /learningtorun_2018