



تاریخ:
شماره:
پیوست:

نام و نام خانوادگی: علی بنی‌اسد شماره دانشجویی: ۴۰۱۲۰۹۲۴۴ معدل: ۱۶/۳۴
گرایش: فضا تعداد واحدهای گذرانده: ۱۵ استاد راهنما: دکتر نوبهارى
استاد راهنمای همکار: تعداد واحد پروژه: ۶ استاد ممتحن:

عنوان کامل پروژه:

فارسی:

هدایت بازی دیفرانسیلی با استفاده از یادگیری تقویتی مقاوم در محیط‌های پویا چندجسمی با پیشران کم
انگلیسی:

Robust Reinforcement Learning Differential Game Guidance in Low-Thrust, Multi-Body Dynamical Environments

نوع پروژه: ■ کاربردی: ■ بنیادی: □ توسعه‌ای: ■

معرفی موضوع:

این پژوهش یک منطق هدایت مقاوم با استفاده از یادگیری تقویتی بازی‌های دیفرانسیلی^۱ را ارائه می‌دهد. این منطق هدایت در بستر شبکه عصبی^۲ برای راهنمایی خودکار و حلقه بسته فضایی با پیشران کم در محیط‌های چند جسمی متغیر مانند منظومه زمین-ماه، ارائه می‌شود.

اهمیت موضوع:

مطالعه و توسعه این الگوریتم هدایت جدید با استفاده از الگوریتم‌های یادگیری تقویتی و یادگیری بازی‌های دیفرانسیلی دارای اهمیت ویژه‌ای است. این الگوریتم هدایت، به عنوان یک مدل مقاوم، قابلیت مقابله با نوسانات و متغیرهای محیطی را داراست. علاوه بر این، ساختار جدید پردازنده‌های کامپیوتر پرواز، با امکانات به‌روز و محاسبات ماتریسی بهبود یافته، اجرای بهینه‌تر الگوریتم‌های یادگیری ماتریسی را آسان کرده و امکان اجرای موثر آنها را فراهم کرده است.

با توجه به اینکه ایستگاه زمینی در بیشتر زمان ماموریت فاصله زمانی قابل توجهی از محل اجرای دستورات و مکان اعمال دستورات دارد، استفاده از این الگوریتم هدایت می‌تواند به عنوان یک تکنولوژی بسیار مهم در کاهش تاخیرهای ناشی از این فواصل زمانی و افزایش پایداری و دقت سیستم‌های کامپیوتر پرواز تلقی شود. توانایی این الگوریتم هدایت در انجام محاسبات پیچیده و تصمیم‌گیری‌های برنامه‌ریزی در زمان واقعی، اهمیت ویژه‌ای برای امنیت و پایداری پروازهای فضایی دارد و می‌تواند تاثیر مثبتی در اجرای موثر ماموریت‌های فضایی آینده داشته باشد.

کاربردها:

کاربردهای الگوریتم هدایت در ماموریت‌های مختلف فضایی در ادامه آورده شده است. این الگوریتم هدایت می‌تواند در ماموریت‌های فضایی مختلفی مانند ماموریت‌های ماهواره‌ای، ماموریت‌های سفر به ماه و ماموریت‌های نزدیک به زمین مورد استفاده قرار گیرد.

۱. هدایت خودکار فضاییها: یکی از کاربردهای اصلی این پژوهش، هدایت خودکار فضاییها در محیط‌های پویا و پیچیده فضایی است. این پژوهش به ایجاد یک سیستم هدایت مبتنی بر یادگیری تقویتی که توانایی انجام مانورها، تعقیب مسیرهای مرجع بین مدارهای مختلف و بهبود اشتباهات را دارد، می‌پردازد.

¹ Differential Game Reinforcement Learning (DGRL)

² Neural Network (NN)



تاریخ:

شماره:

پیوست:

۲. حل مسائل پیچیده هدایت: این پژوهش برای حل مسائل پیچیده هدایت فضایی که نیازمند تصمیم‌گیری در زمان واقعی و مطابق با متغیرهای محیطی می‌باشند، استفاده می‌شود. از جمله این مسائل می‌توان به تغییر مسیر در محیط‌های پیچیده، تنظیمات نیرویه پیشران، و اصلاح وضعیت نسبت به مدارهای مرجع اشاره کرد.

۳. هدایت در محیط‌های نامطمئن: از طریق استفاده از الگوریتم‌های یادگیری تقویتی، این پژوهش به هدایت در محیط‌های نامطمئن و بدون نیاز به مدل دقیق محیط فضایی امکان می‌دهد. این موضوع در مواقعی که مدل‌سازی دقیق مدل‌های دینامیکی مشکل باشد، قابل اجرا است.

۴. استفاده در ماموریت‌های متنوع: این پژوهش امکان استفاده در ماموریت‌های مختلف فضایی را فراهم می‌کند. از جمله ماموریت‌های مختلف می‌توان به ماموریت‌های ماهواره‌ای، تعمیر و نگهداری ماموریت فضایی، تعقیب اهداف در مدارهای مختلف، و انجام مانورهای پیچیده اشاره کرد.

تعریف دقیق مسئله:

در پژوهش حاضر، مسأله هدایت فضاپیما برای سیستم زمین-ماه با استفاده از مدل دینامیکی سه‌جسم محدود دایره‌ای^۳ مطرح می‌شود. در این مسأله، یک فضاپیما با سیستم پیشران کم به یک مسیر مرجع در سیستم CR3BP منتقل می‌شود. شرایط اولیه فضاپیما از مسیر مرجع انحراف‌های تصادفی دارند. هدف از این پژوهش، توسعه یک الگوریتم هدایت حلقه بسته با استفاده از اصول یادگیری تقویتی بر مبنای بازی دیفرانسیلی است که به فضاپیما اجازه می‌دهد با فرض بدترین اغتشاش به محیط مرجع بازگردد، مسیر مرجع را دنبال کرده و به مدار مقصد برسد. بازی دیفرانسیلی موجب می‌شود الگوریتم هدایت نه تنها بهینه باشد بلکه مقاوم نیز باشد.

فرضیات مسئله:

- مدل دینامیکی مسأله، مسأله سه‌جسم محدود دایره‌ای سیستم زمین-ماه است. این فرضیه یک سیستم دوجسمی ساده با فضاپیما به عنوان جسم سوم بسیار کوچک را فرض می‌کند.
- حرکت فضاپیما تنها در صفحه اتفاق می‌افتد و تنها نیروهای گرانشی در نظر گرفته می‌شوند. سایر انحراف‌های مانند فشار تابشی خورشید نادیده گرفته می‌شوند.
- فضاپیما به یک سیستم پیشران کم مجهز است که شعاع تاثیر مشخصی دارد. میزان و جهت پیشران به صورت مستمر تغییر کرده و قابل تنظیم است.
- مسأله هدایت شامل انتقال بین مدارهای دوره‌ای (مدارهای لیپانوف) در مدل CR3BP زمین-ماه است.
- توانایی موتور دارای محدودیت است. موتور قادر به تولید پیشران در هر جهت و مقدار نیست.
- وضعیت فضاپیما به طور کامل قابل مشاهده و در دسترس الگوریتم هدایت در هر مرحله زمانی است.
- محیط یادگیری تقویتی ویژگی‌های مارکوف را دارد، به این معنا که وضعیت فعلی تمام اطلاعات لازم برای پیش‌بینی وضعیت‌های آینده را فراهم می‌کند.

³Circular Restricted Three-Body Problem (CR3BP)



تاریخ:
شماره:
پیوست:

- آموزش در یک محیط چرخه‌ای^۴ انجام می‌شود.
- تابع پاداش به صورت بازی دیفرانسیلی طراحی شده است که فضاپیما در مسیر نزدیک به مرجع با فرض بدترین اغتشاش بماند و به مدار مقصد برسد.
- هدف اصلی تنها اجرا پذیری مسأله نیست و دیگر اهداف مانند مصرف کمتر سوخت یا زمان انتقال با استفاده از نظریه بازی در نظر گرفته می‌شوند.
- این رویکرد نیاز به دسترسی به منابع محاسباتی با عملکرد بالا برای آموزش یادگیری تقویتی قبل از اجرای آن در رایانه پرواز فضاپیما را دارد.

روش انجام کار:

• مرحله ۱: ایجاد محیط دینامیکی

- پیاده‌سازی مدل دینامیکی محیط مورد نظر انجام می‌شود. این مدل باید شامل تمام جوانب مهم محیط از جمله قوانین حرکت، ابعاد فضایی، و وضعیت اولیه باشد.
- در صورت پیچیدگی مدل دینامیکی محیط، از کتابخانه‌ها و ابزارهای مهندسی نرم‌افزاری مانند NumPy و SciPy در پایتون برای پیاده‌سازی آن استفاده می‌شود.
- مدل دینامیکی محیط با ویژگی اینک به درستی کار کند و اطلاعات مورد نیاز برای یادگیری تقویتی (مانند وضعیت فعلی) در دسترس باشند، ایجاد می‌شود.

• مرحله ۲: ایجاد بازیگر هدایت

- در این مرحله، بازیگری برای انجام هدایت در محیط خود ایجاد می‌شود. این بازیگر مسئول انتخاب اعمال (اعمال کم پیشران) بر اساس وضعیت فعلی محیط است.
- از الگوریتم‌های یادگیری تقویتی مختلف مانند Q-Learning، Deep Q-Networks (DQN)، یا Proximal Policy Optimization (PPO) برای ایجاد بازیگر هدایت استفاده می‌شود. انتخاب الگوریتم بستگی به محیط و اهداف آموزش دارد.

• مرحله ۳: استفاده از منابع GPU در Google Colab

- Google Colab به عنوان یک محیط آموزش آنلاین به کار می‌رود که اجازه می‌دهد کدهای پایتون در یک محیط مبتنی بر مرورگر اجرا شوند. از این مزیت این محیط می‌توان به دسترسی به منابع GPU برای آموزش مدل‌های عمیق اشاره کرد.
- برای استفاده از GPU در Google Colab، می‌توان از کتابخانه‌های معروفی مانند TensorFlow یا PyTorch که از GPU پشتیبانی می‌کنند، استفاده می‌شود. به این ترتیب، آموزش مدل‌های یادگیری تقویتی بسیار سریع‌تر و رایگان انجام می‌شود.

پیشینه‌ی موضوع:

⁴episodic



فرم تعریف پروژه کارشناسی ارشد

تاریخ:

شماره:

پیوست:

رایج‌ترین تکنیک‌های هدایت معمولاً با استفاده از ایستگاه‌های کنترل مستقر در زمین انجام می‌شوند. با این حال، حساسیت به شکست ارتباطات، تاخیرهای زمانی، محدودیت‌های انتقال داده، پیچیدگی سنسورها و هزینه‌های عملیاتی همگی از دلایلی هستند که هدایت، مسیریابی و کنترل^۵ را از کامپیوتر ایستگاه زمینی به کامپیوتر پرواز منتقل می‌کنند. در حالی که در طراحی مسیر از افزایش قدرت سخت‌افزارهای کامپیوتر بهره‌برداری می‌کنند، تعداد کمی از آن‌ها برای اجرای خودکار درون‌سفینه عملی هستند. این به دلیل محدودیت‌های منابع محاسباتی است که در فضاپیماها وجود دارد.

در فرآیند طراحی مسیر، معمولاً یک مسیر بهینه و تاریخچه فرمان کنترلی طراحی می‌شود که با معیارهای مأموریت، مصرف سوخت و زمان پرواز، مطابقت داشته باشد. این روش قبل از پرواز انجام می‌شود و می‌تواند از استراتژی‌های متعددی برای هدایت بهینه با پیشران کم استفاده کند، از جمله تکنیک‌های بهینه‌سازی جهانی [۱] و برنامه‌نویسی غیرخطی [۲]. با این حال، توانایی سریع محاسبه مجدد مسیر مرجع و تاریخچه فرمان کنترلی در داخل فضاپیما در هنگام پرواز بسیار مهم است. برخی از پژوهش‌ها تلاش می‌کنند که روش‌های بهینه‌سازی را در کامپیوتر سفینه اجرا کنند، برای مثال از دینامیک‌های ساده‌تری که جسم سوم در آن نیست، استفاده می‌شود. با نگاه به هدایت درون‌سفینه‌ای از دیدگاه یادگیری ماشین، یک کنترل‌کننده شبکه عصبی حلقه بسته، امکان محاسبه سریع و خودکار تاریخچه کنترل را برای یک فضاپیما فراهم می‌کند. به علاوه، یادگیری تقویتی^۶ یادگیری پیش از پرواز واقعی است و می‌تواند از سخت‌افزارهای سریع و قدرتمند و ارزان قیمت زمینی جهت یادگیری استفاده کند. روش‌های هدایت و بهینه‌سازی مسیر فضاپیماها به طور کلی به راه‌حل‌های اولیه مناسب نیاز دارند. در مسائل چندجسمی، طراحان مسیر اغلب حدس‌های اولیه کم‌هزینه‌ای برای انتقال‌ها با استفاده از نظریه سیستم‌های دینامیکی و صفحات نامتعوض [۳، ۴] ایجاد می‌کنند. روش‌های مبتنی بر سیستم‌های دینامیکی در بسیاری از کاربردهای گذشته مفید بوده‌اند و در ترکیب با اصلاحات دیفرانسیل و یا تکنیک‌های بهینه‌سازی، برای بسیاری از کاربردها راه‌حل‌های بهینه تولید می‌کنند. با این حال، این رویکرد محاسباتی، فشرده است و اغلب به تعاملات انسان در حلقه نیاز دارد. به عنوان یک جایگزین، تکنیک‌های بهینه‌سازی جهانی ابتکاری مانند جستجوی حوضچه و الگوریتم‌های تکاملی، نیاز به راه‌حل‌های راه‌اندازی دقیق را کاهش می‌دهند [۱]، اما پیچیدگی محاسباتی مربوطه آنها را برای استفاده درون‌سفینه ناممکن می‌سازد.

شبکه‌های عصبی ویژگی‌های جذابی را برای فعال‌سازی automatic guidance در فضاپیما دارند. به عنوان مثال، شبکه‌های عصبی می‌توانند به‌طور مستقیم از تخمین‌های وضعیت به دستورهای پیشران کنترلی که با محدودیت‌های مأموریت سازگار است، برسند. عملکرد هدایت شبکه‌های عصبی در مطالعاتی مانند فرود بر سیارات [۲۶]، عملیات نزدیکی به سیارات [۲۵] و کنترل فضاپیما با پیشران ازدست رفته [۱۳] نشان داده شده است. تازه‌ترین پیشرفت‌های تکنیک‌های یادگیری ماشینی در مسائل onboard automation به‌طور گسترده‌ای مورد مطالعه قرار گرفته‌اند، از پژوهش‌های اولیه تا توانایی‌های پیاده‌سازی. به عنوان مثال، الگوریتم‌های یادگیری ماشینی ابتدایی در فضاپیماهای مریخی برای کمک به شناسایی ویژگی‌های زمین‌شناسی تعبیه شده‌اند. الگوریتم AEGIS توانایی انتخاب خودکار هدف ChemCam در داخل فضاپیماهای Spirit، Opportunity و Curiosity را فعال می‌کند [۵]. در کامپیوتر پرواز اصلی Curiosity، فرآیند بهبود نقطه‌گیری^۷ نیاز به ۹۴ تا ۹۶ ثانیه دارد [۶]، که به طور قابل توجهی کمتر از زمان مورد نیاز برای ارسال تصاویر به زمین و انتظار برای انتخاب دستی توسط دانشمندان است. برنامه‌های آینده برای کاربردهای یادگیری ماشینی درون‌سفینه شامل توانایی‌های رباتیکی درون‌سفینه برای فضاپیمای Perseverance [۸، ۹] و شناسایی عیب برای Europa Clipper [۷] می‌شوند. الگوریتم‌های یادگیری ماشینی پتانسیلی برای سهم مهمی در مأموریت‌های اتوماسیون آینده دارند.

علاوه بر رباتیک سیاره‌ای، تحقیقات مختلفی به تکنیک‌های مختلف یادگیری ماشینی برای مسائل astrodynamics اختصاص داده‌اند. در طراحی مسیر، وظایف رگرسیون معمولاً مؤثرتر هستند. به عنوان مثال، Dachwald در سال ۲۰۰۴ از یک شبکه عصبی کم‌عمق (NN)

^۵Guidance, Navigation and Control (GNC)

^۶Reinforcement Learning (RL)

^۷(refinement process)



بسمه تعالی

دانشکده مهندسی هوافضا

فرم تعریف پروژه کارشناسی ارشد

تاریخ:

شماره:

پیوست:

در بهینه‌سازی مسیرهای رانشگر کم پیشران استفاده کرد [۱۰]. تحقیقات جدید شامل شناسایی انتقال‌های هتروکلینیک [۱۱]، اصلاح مسیر رانشگر کم پیشران [۱۲] و تجزیه و تحلیل مشکلات ازدست رفتن رانشگر [۱۳] می‌شوند. تکنیک‌های یادگیری نظارتی می‌توانند نتایج مطلوبی تولید کنند، اما دارای محدودیت‌های قابل توجهی هستند. ابتدا، این رویکردها بر وجود دانش از پیش از فرآیند تصمیم‌گیری متکی هستند. کاربر با انتخاب نتایج مطلوب، فرض می‌کند که این دانش را دارد. این امر مستلزم دقیق بودن داده‌های تولیدشده توسط کاربر برای نتایج مطلوب و همچنین وجود تکنیک‌های موجود برای حل مشکل کنونی و تولید داده است. در بخش‌هایی که چنین دانشی وجود ندارد، تکنیک‌های یادگیری نظارتی قابل استفاده نیستند. در سال‌های اخیر، یادگیری تقویتی (RL) در دستیابی به عملکرد بهینه در دامنه‌هایی با ابهام محیطی قابل توجه، به اثبات رسیده است [۱۴، ۱۵]. هدایت انجام‌شده توسط RL را می‌توان به صورت گسترده بر اساس فاز پرواز دسته‌بندی کرد. مسائل فرود [۱۶، ۱۷] و عملیات در نزدیکی اجسام کوچک [۲۵، ۲۶]، از حوزه‌های پژوهشی هستند که از RL استفاده می‌کنند. تحقیقات دیگر شامل مواجهه تداخل خارجی جوی [۱۹]، نگهداری ایستگاهی [۲۰] و جلوگیری از شناسایی [۲۱] است. مطالعاتی که فضایی‌های رانشگر کم پیشران را در یک چارچوب دینامیکی چندبدنی با استفاده از RL انجام شده است شامل طراحی انتقال با استفاده از Q-learning [۲۲]، Proximal Policy Optimization [۲۳] و هدایت نزدیکی مدار [۲۴] است.

اهداف پژوهش:

- طراحی یک بازیکن جهت هدایت به کمک یادگیری تقویتی
- طراحی یک الگوریتم هدایت مقاوم به کمک بازی دیفرانسیلی
- بهبود عملکرد فضایی‌ها در محیط سه جسم
- عدم نیاز به ایستگاه زمینی جهت صدور فرمان هدایت
- پیاده‌سازی الگوریتم درون‌سفینه‌ای در محیط شبیه‌سازی شده



بسمه تعالی

دانشکده مهندسی هوافضا

فرم تعریف پروژه کارشناسی ارشد

تاریخ:

شماره:

پیوست:

جدول ۱: جدول برنامه‌ریزی پروژه نوابری اینرسی مشارکتی مبتنی بر هوش مصنوعی

فصل	فعالیت‌ها
تابستان سال اول	انجام مطالعه و مروری جامع در زمینه بازی دیفرانسیلی و یادگیری تقویتی.
	بررسی و تعریف دقیق مسأله یادگیری تقویتی و الگوریتم هدایت به وسیله یادگیری تقویتی در فضای سه‌جسم.
	طراحی ساختار محیط سه‌جسم و بازیکن یادگیری تقویتی و ارتباط بین محیط و بازیکن.
نیمسال اول سال دوم	طراحی مسیر انتقالی بهینه بین دو مدار
	شبیه‌سازی دقیق محیط مسأله سه‌جسم جهت آموزش بازیکن یادگیری تقویتی
	طراحی و توسعه سیستم یادگیری تقویتی جهت محاسبه دستور
نیمسال دوم سال دوم	پیاده‌سازی هدایت فضایی با استفاده از الگوریتم‌های کلاسیک
	آموزش الگوریتم‌های یادگیری تقویتی در محیط‌های شبیه‌سازی شده
تابستان سال دوم	اصلاح و بهبود عملکرد یادگیری تقویتی در هدایت فضایی
	نتیجه‌گیری از تحقیقات و ارائه گزارش نهایی شامل معرفی مسئله، روش‌های استفاده‌شده، نتایج حاصل و پیشنهادات برای تحقیقات آتی

نتایج مورد انتظار:



بسمه تعالی

دانشکده مهندسی هوافضا

فرم تعریف پروژه کارشناسی ارشد

تاریخ:

شماره:

پیوست:

- هدایت فضایی بدون نیاز به ایستگاه زمینی
- کاهش هزینه ساخت به دلیل عدم نیاز به سیستم‌های ارتباطی قوی
- افزایش ایمنی مأموریت به دلیل استفاده از بازی دیفرانسیلی
- کاهش مصرف سوخت و هزینه به دلیل بهینه بودن الگوریتم هدایت

روش صحت‌سنجی نتایج:

- مقایسه با سایر روشهای معتبر
- مدلسازی و شبیه‌سازی سیستم محیط و بازیگر

گلوگاه‌های پیش‌بینی‌شده:

- آموزش شبکه: به علت پیچیده بودن شبکه و محطی که طراحی شده است، محاسبات سنگین می‌شود.
- پایداری و کنترل: به علت غیرخطی بودن دینامیک سیستم و ناپایداری ذاتی سیستم، کنترل و پایداری آن مشکل است.

نوآوری‌ها:

- طراحی الگوریتم هدایت بهینه درون‌سفینه‌ای
- طراحی الگوریتم هدایت مقاوم درون‌سفینه‌ای
- طراحی الگوریتم هدایت یادگیری تقویتی بر مبنای بازی دیفرانسیلی

References

- [1] Vavrina, Matthew A, Jacob A Englander, Sean M Phillips, and Kyle M Hughes. "Global, multi-objective trajectory optimization with parametric spreading." In AAS AIAA Astrodynamics Specialist Conference 2017, Tech. No. GSFC-E-DAA-TN45282, 2017.
- [2] Ocampo, Cesar. "Finite Burn Maneuver Modeling for a Generalized Spacecraft Trajectory Design and Optimization System." Annals of the New York Academy of Sciences 1017 (2004): 210-233. doi: 10.1196/annals.1311.013.



تاریخ:

شماره:

پیوست:

- [3] Haapala, A. F., & Howell, K. C. (2016). "A framework for constructing transfers linking periodic libration point orbits in the spatial circular restricted three-body problem." *International Journal of Bifurcation and Chaos*, 26(05), 1630013.
- [4] Marchand, B. G. and Scarritt, S. K. and Pavlak, T. A. and Howell, K. C. "A dynamical approach to precision entry in multi-body regimes: Dispersion manifolds." *Acta Astronautica* 89 (2013): 107-120. doi: 10.1016/j.actaastro.2013.02.015.
- [5] T.A. Estlin, B.J. Bornstein, D.M. Gaines, R.C. Anderson, D.R. Thompson, M. Burl, R. Castaño, M. Judd. "Aegis automated science targeting for the MER Opportunity rover." *ACM Trans. Intell. Syst. Technol. (TIST)*, 3, 1-19, 2012.
- [6] R. Francis, T. Estlin, G. Doran, S. Johnstone, D. Gaines, V. Verma, M. Burl, J. Frydenvang, S. Montano, R. Wiens, S. Schaffer, O. Gasnault, L. Deflores, D. Blaney, B. Bornstein. "Aegis autonomous targeting for ChemCam on Mars Science Laboratory: Deployment and results of initial science team use." *Science Robotics*, 2, 2017.
- [7] K.L. Wagstaff, G. Doran, A. Davies, S. Anwar, S. Chakraborty, M. Cameron, I. Daubar, C. Phillips. "Enabling onboard detection of events of scientific interest for the Europa Clipper spacecraft," in: 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Association for Computing Machinery, Anchorage, Alaska, 2019, pp. 2191–2201, doi: 10.1145/3292500.3330656.
- [8] S. Higa, Y. Iwashita, K. Otsu, M. Ono, O. Lamarre, A. Didier, M. Hoffmann. "Vision-based estimation of driving energy for planetary rovers using deep learning and terramechanics," *IEEE Robot. Autom. Lett.* 4 (2019) 3876–3883.
- [9] B. Rothrock, J. Papon, R. Kennedy, M. Ono, M. Heverly, C. Cunningham. "Spoc: Deep learning-based terrain classification for Mars rover missions," in: AIAA Space and Astronautics Forum and Exposition, SPACE 2016, American Institute of Aeronautics and Astronautics Inc, AIAA, 2016, pp. 1–12.
- [10] B. Dachwald. "Evolutionary neurocontrol: A smart method for global optimization of low-thrust trajectories," in: AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Providence, Rhode Island, 2004, pp. 1–16.
- [11] S. De Smet, D.J. Scheeres. "Identifying heteroclinic connections using artificial neural networks," *Acta Astronaut.* 161 (2019) 192–199.
- [12] N.L.O. Parrish. "Low Thrust Trajectory Optimization in Cislunar and Translunar Space (Ph.D. thesis)," University of Colorado Boulder, 2018.
- [13] A. Rubinsztein, R. Sood, F.E. Laipert. "Neural network optimal control in astrodynamics: Application to the missed thrust problem," *Acta Astronaut.* 176 (2020) 192–203.
- [14] N. Heess, D. TB, S. Sriram, J. Lemmon, J. Merel, G. Wayne, Y. Tassa, T. Erez, Z. Wang, S.M.A. Eslami, M.A. Riedmiller, D. Silver. "Emergence of locomotion behaviours in rich environments," 2017. CoRR abs/1707.02286. arXiv:1707.02286.
- [15] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, D. Hassabis. "Mastering the game of Go without human knowledge," *Nature* 550 (2017) Article.



..... تاریخ:

..... شماره:

..... پیوست:

-
- [16] R. Furfaro, A. Scorsoglio, R. Linares, M. Massari, "Adaptive generalized ZEM-ZEV feedback guidance for planetary landing via a deep reinforcement learning approach," *Acta Astronaut.* 171 (2020) 156–171.
- [17] B. Gaudet, R. Linares, R. Furfaro, "Deep reinforcement learning for six degrees of freedom planetary landing," *Adv. Space Res.* 65 (2020) 1723–1741.
- [18] J. Broida, R. Linares, "Spacecraft rendezvous guidance in cluttered environments via reinforcement learning," in: *29th AAS/AIAA Space Flight Mechanics Meeting*, American Astronautical Society, Ka'anapali, Hawaii, 2019, pp. 1–15.
- [19] B. Gaudet, R. Furfaro, R. Linares, "Reinforcement learning for angle-only intercept guidance of maneuvering targets," *Aerosp. Sci. Technol.* 99 (2020).
- [20] D. Guzzetti, "Reinforcement learning and topology of orbit manifolds for station-keeping of unstable symmetric periodic orbits," in: *AAS/AIAA Astrodynamics Specialist Conference*, American Astronautical Society, Portland, Maine, 2019, pp. 1–20.
- [21] J.A. Reiter, D.B. Spencer, "Augmenting spacecraft maneuver strategy optimization for detection avoidance with competitive coevolution," in: *20th AIAA Scitech Forum*, AIAA, Orlando, Florida, 2020, pp. 1–11.
- [22] A. Das-Stuart, K.C. Howell, D.C. Folta, "Rapid trajectory design in complex environments enabled by reinforcement learning and graph search strategies," *Acta Astronaut.* 171 (2020) 172–195.
- [23] D. Miller, R. Linares, "Low-thrust optimal control via reinforcement learning," in: *29th AAS/AIAA Space Flight Mechanics Meeting*, American Astronautical Society, Kaanapali, Hawaii, 2019, pp. 1–18.
- [24] C.J. Sullivan, N. Bosanac, "Using reinforcement learning to design a low-thrust approach into a periodic orbit in a multi-body system," in: *20th AIAA Scitech Forum*, AIAA, Orlando, Florida, 2020, pp. 1–19.
- [25] B. Gaudet, R. Linares, R. Furfaro, "Terminal adaptive guidance via reinforcement meta-learning: Applications to autonomous asteroid close-proximity operations," *Acta Astronaut.* 171 (2020) 1–13.
- [26] B. Gaudet, R. Linares, R. Furfaro, "Six degree-of-freedom hovering over an asteroid with unknown environmental dynamics via reinforcement learning," in: *20th AIAA Scitech Forum*, AIAA, Orlando, Florida, 2020, pp. 1–15.