



## فرم تعریف پروژه کارشناسی ارشد

تاریخ: ۱۴۰۲/۰۷/۲۶  
 شماره: .....  
 پیوست: .....

نام و نام خانوادگی: علی بنی‌اسد شماره دانشجویی: ۴۰۱۲۰۹۲۴۴ معدل: ۱۶/۳۴  
 گرایش: فضا تعداد واحدهای گذرانده: ۱۵ استاد راهنما: دکتر نوبهاری  
 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:

در سال‌های اخیر، پیشرفت‌های فناوری در زمینه‌های مختلف، از جمله کنترل پرواز، پردازش سیگنال و هوش مصنوعی، به افزایش کاربردهای ماهواره با پیشران کم در منظومه زمین-ماه کمک کرده است. ماهواره با پیشران کم می‌تواند برای تعقیب ماهواره‌ها، انتقال مداری و استقرار ماهواره‌ها استفاده شود. این فناوری می‌تواند به کاهش هزینه‌ها، بهبود عملکرد و کاهش خطرات مرتبط به پروازهای فضایی کمک کند. این پژوهش یک منطق هدایت مقاوم با استفاده از یادگیری تقویتی بازی‌های دیفرانسیلی<sup>۱</sup> را ارائه می‌دهد. این منطق هدایت در بستر شبکه عصبی<sup>۲</sup> برای راهنمایی خودکار و حلقه بسته فضایی‌هایی 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 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.

<sup>1</sup> Differential Game Reinforcement Learning (DGRL)<sup>2</sup> Neural Network (NN)



## فرم تعریف پروژه کارشناسی ارشد

تاریخ: ۱۴۰۲/۰۷/۲۶  
 شماره: .....  
 پیوست: .....

۱. هدایت خودکار فضاپیماها: یکی از کاربردهای اصلی این پژوهش، هدایت خودکار فضاپیماها 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.

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<sup>۳</sup> 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 لیاپانوف) در مدل CR3BP زمین-ماه است.
- توانایی موتور دارای محدودیت است. موتور قادر به تولید پیشران در هر جهت و مقدار نیست.

<sup>3</sup>Circular Restricted Three-Body Problem (CR3BP)



تاریخ: ۱۴۰۲/۰۷/۲۶

شماره: .....

پیوست: .....

- 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.
- آموزش در یک محیط چرخه‌ای<sup>۴</sup> 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 اینکه 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.

<sup>4</sup>episodic



## فرم تعریف پروژه کارشناسی ارشد

تاریخ: ۱۴۰۲/۰۷/۲۶

شماره: .....

پیوست: .....

– 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:

#### پژوهش‌های خارجی:

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<sup>۵</sup> 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<sup>۶</sup> 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 نامتعین [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<sup>۷</sup> نیاز به ۹۴ تا ۹۶ ثانیه دارد [۶]، که به طور قابل توجهی کمتر از زمان مورد نیاز برای ارسال تصاویر به زمین

<sup>۵</sup>Guidance, Navigation and Control (GNC)

<sup>۶</sup>Reinforcement Learning (RL)

<sup>۷</sup>(refinement process)



## فرم تعریف پروژه کارشناسی ارشد

تاریخ: ۱۴۰۲/۰۷/۲۶  
 شماره: .....  
 پیوست: .....

و انتظار برای انتخاب دستی توسط دانشمندان است. برنامه‌های آینده برای کاربردهای یادگیری ماشینی درون‌سفینه شامل توانایی‌های رباتیکی درون‌سفینه برای فضایی‌های 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) 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.

### پژوهش‌های داخلی:

ناوبری و کنترل فضایی‌ها و ماهواره‌ها، یکی از چالش‌های مهم در مهندسی فضایی است. این مسئله شامل تعیین موقعیت و جهت فضایی‌ها در فضا و هدایت و کنترل حرکت آن برای رسیدن به اهداف مورد نظر است. در سال‌های اخیر، پیشرفت‌های زیادی در زمینه ناوبری و کنترل فضایی‌ها و ماهواره‌ها حاصل شده است. این پیشرفت‌ها عمدتاً ناشی از توسعه روش‌های جدیدی مانند یادگیری تقویتی، کنترل‌های تطبیقی و بینایی کامپیوتری است. به عنوان مثال، در [۲۷]، یک روش جدید برای انجام مانور اتصال دو فضایی‌ها با استفاده از الگوریتم یادگیری تقویتی کیو ارائه شده است. این روش می‌تواند با دقت بسیار بالایی، مانور اتصال دو فضایی‌ها را انجام دهد. در [۲۸]، یک کنترل‌کننده تطبیقی بهینه بدون مدل برای بازپیکربندی آرایش پروازی ماهواره‌ها ارائه شده است. این کنترل‌کننده می‌تواند برای انواع مختلف آرایش پروازی ماهواره‌ها استفاده شود. در [۲۹]، یک روش جدید برای ناوبری و کنترل بازوی رباتیک ماهواره پرواز-آزاد ارائه شده است. این روش بر مبنای یک سیستم مبتنی بر بینایی کامپیوتری طراحی شده است. پیشرفت‌های اخیر در زمینه ناوبری و کنترل فضایی‌ها و ماهواره‌ها، به بهبود ایمنی و کارایی مأموریت‌های فضایی کمک کرده 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



تاریخ: ۱۴۰۲/۰۷/۲۶

شماره: .....

پیوست: .....

- پیاده‌سازی الگوریتم درون‌سفینه‌ای در محیط شبیه‌سازی شده

جدول ۱: جدول برنامه‌ریزی پروژه ناولبری 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 فضاپیما با استفاده از الگوریتم‌های کلاسیک
	آموزش الگوریتم‌های یادگیری تقویتی در محیط‌های شبیه‌سازی شده
تابستان سال دوم	اصلاح و بهبود عملکرد یادگیری تقویتی در هدایت فضاپیما
	نتیجه‌گیری از تحقیقات و ارائه گزارش نهایی شامل معرفی مسئله، روش‌های استفاده‌شده، نتایج حاصل و پیشنهادات برای تحقیقات آتی

نتایج مورد انتظار:

- هدایت فضاپیما 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:



تاریخ: ۱۴۰۲/۰۷/۲۶

شماره: .....

پیوست: .....

- 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



تاریخ: ۱۴۰۲/۰۷/۲۶

شماره: .....

پیوست: .....

## مراجع

- [۱] multi- "Global, Hughes. M Kyle and Phillips, M Sean Englander, A Jacob A, Matthew Vavrina, Specialist Astrodynamics AIAA AAS In spreading." parametric with optimization trajectory objective .۲۰۱۷ GSFC-E-DAA-TN۴۵۲۸۲, No. Tech. ,۲۰۱۷ Conference
- [۲] Design Trajectory Spacecraft Generalized a for Modeling Maneuver Burn "Finite Cesar. Ocampo, Optimization and the of Annals System." ۱۰۱۷ Sciences of Academy York New : (۲۰۰۴) ۲۱۰-۲۳۳. doi: ۱۳.۱۳۱۱.annals/۱۱۹۶.۱۰
- [۳] periodic linking transfers constructing for framework "A .(۲۰۱۶) C. K. Howell, & F., A. Haapala, of Journal International problem." three-body restricted circular spatial the in orbits point libration .۱۶۳۰۰۱۳, (۰۵)۲۶ Chaos, and Bifurcation
- [۴] to approach dynamical "A C. K. Howell, and A. T. Pavlak, and K. S. Scarritt, and G. B. Marchand, Dispersion regimes: multi-body in entry precision -۱۰۷ : (۲۰۱۳) ۸۹ Astronautica Acta manifolds." .۱۵.۰۲.۲۰۱۳.j.actaastro/۱۰۱۶.۱۰ doi: .۱۲۰
- [۵] M. Castaño, R. Burl, M. Thompson, D.R. Anderson, R.C. Gaines, D.M. Bornstein, B.J. Estlin, T.A. Syst. Intell. Trans. ACM rover." Opportunity MER the for targeting science automated "Aegis Judd. (TIST), Technol. ۲۰۱۲, ۱۹-۱, ۳
- [۶] Mon- S. Frydenvang, J. Burl, M. Verma, V. Gaines, D. Johnstone, S. Doran, G. Estlin, T. Francis, R. autonomous "Aegis Bornstein. B. Blaney, D. Deflores, L. Gasnault, O. Schaffer, S. Wiens, R. tano, team science initial of results and Deployment Laboratory: Science Mars on ChemCam for targeting .۲۰۱۷, ۲ Robotics, Science use."
- [۷] Phillips. C. Daubar, I. Cameron, M. Chakraborty, S. Anwar, S. Davies, A. Doran, G. Wagstaff, K.L. ۲۵th in: spacecraft." Clipper Europa the for interest scientific of events of detection onboard "Enabling for Association Mining, Data & Discovery Knowledge on Conference International SIGKDD ACM .۳۳۳۰۶۵۶.۱۱۴۵/۳۲۹۲۵۰۰.۱۰ doi: ,۲۲۰۱-۲۱۹۱ pp. ,۲۰۱۹ Alaska, Anchorage, Machinery, Computing
- [۸] esti- "Vision-based Hoffmann. M. Didier, A. Lamarre, O. Ono, M. Otsu, K. Iwashita, Y. Higa, S. Robot. IEEE terramechanics," and learning deep using rovers planetary for energy driving of mation .۳۸۸۳-۳۸۷۶ (۲۰۱۹) ۴ Lett. Autom.
- [۹] learning- Deep "Spoc: Cunningham. C. Heverly, M. Ono, M. Kennedy, R. Papon, J. Rothrock, B. and Forum Astronautics and Space AIAA in: missions," rover Mars for classification terrain based pp. ,۲۰۱۶ AIAA, Inc. Astronautics and Aeronautics of Institute American ,۲۰۱۶ SPACE Exposition. .۱۲-۱





تاریخ: ۱۴۰۲/۰۷/۲۶  
شماره: .....  
پیوست: .....

- low-thrust of optimization global for method smart A neurocontrol: "Evolutionary Dachwald. B. [۱۰]  
Rhode Providence, Exhibit, and Conference Specialist Astrodynamics AIAA/AAS in: trajectories."  
Island, ۲۰۰۴, pp. ۱-۱۶.
- networks," neural artificial using connections heteroclinic "Identifying Scheeres. D.J. Smet, De S. [۱۱]  
Astronaut. Acta ۱۶۱ (۲۰۱۹) ۱۹۲-۱۹۹.
- the- (Ph.D. Space Translunar and Cislunar in Optimization Trajectory Thrust "Low Parrish. N.L.O. [۱۲]  
Boulder, Colorado of University sis), ۲۰۱۸.
- Applica- astrodynamics: in control optimal network "Neural Laipert. F.E. Sood, R. Rubinsztein, A. [۱۳]  
Astronaut. Acta problem," thrust missed the to tion ۱۷۶ (۲۰۲۰) ۱۹۲-۲۰۳.
- S.M.A. Wang, Z. Erez, T. Tassa, Y. Wayne, G. Merel, J. Lemmon, J. Sriram, S. TB, D. Heess, N. [۱۴]  
environments," rich in behaviours locomotion of "Emergence Silver. D. Riedmiller, M.A. Eslami,  
arXiv:۱۷۰۷.۰۲۲۸۶. abs/۱۷۰۷.۰۲۲۸۶. CoRR. ۲۰۱۷
- Baker, L. Hubert, T. Guez, A. Huang, A. Antonoglou, I. Simonyan, K. Schrittwieser, J. Silver, D. [۱۵]  
D. Graepel, T. Driessche, den van G. Sifre, L. Hui, F. Lillicrap, T. Chen, Y. Bolton, A. Lai, M.  
Article. (۲۰۱۷) ۵۵۰ Nature knowledge," human without Go of game the "Mastering Hassabis.
- feedback ZEM-ZEV generalized "Adaptive Massari, M. Linares, R. Scorsoglio, A. Furfaro, R. [۱۶]  
Astronaut. Acta approach," learning reinforcement deep a via landing planetary for guidance  
۱۷۱-۱۵۶ (۲۰۲۰).
- planetary freedom of degrees six for learning reinforcement "Deep Furfaro, R. Linares, R. Gaudet, B. [۱۷]  
Res. Space Adv. landing," ۶۵ (۲۰۲۰) ۱۷۲۳-۱۷۴۱.
- reinforcement via environments cluttered in guidance rendezvous "Spacecraft Linares, R. Broida, J. [۱۸]  
Society, Astronautical American, Meeting Mechanics Flight Space AAS/AIAA ۲۹th in: learning,"  
Hawaii, Ka'anapali, ۲۰۱۹, pp. ۱-۱۵.
- of guidance intercept angle-only for learning "Reinforcement Linares, R. Furfaro, R. Gaudet, B. [۱۹]  
Technol. Sci. Aerosp. targets," maneuvering ۹۹ (۲۰۲۰).
- un- of station-keeping for manifolds orbit of topology and learning "Reinforcement Guzzetti, D. [۲۰]  
American, Conference Specialist Astrodynamics AAS/AIAA in: orbits," periodic symmetric stable  
Maine, Portland, Society, Astronautical ۲۰۱۹, pp. ۱-۲۰.
- detection for optimization strategy maneuver spacecraft "Augmenting Spencer, D.B. Reiter, J.A. [۲۱]  
Florida, Orlando, AIAA, Forum Scitech AIAA ۲۰th in: coevolution," competitive with avoidance  
pp. ۱-۱۱, ۲۰۲۰.



تاریخ: ۱۴۰۲/۰۷/۲۶  
شماره: .....  
پیوست: .....

enabled environments complex in design trajectory "Rapid Foltá, D.C. Howell, K.C. Das-Stuart, A. [۲۲]  
Astronaut. Acta strategies," search graph and learning reinforcement by ۱۷۱ (۲۰۲۰) ۱۹۵-۱۷۲.

AAS/AIAA ۲۹th in: learning," reinforcement via control optimal "Low-thrust Linares, R. Miller, D. [۲۳]  
Hawaii, Kaanapali, Society, Astronautical American, Meeting Mechanics Flight Space ۲۰۱۹, pp. ۱-۱۸.

a into approach low-thrust a design to learning reinforcement "Using Bosanac, N. Sullivan, C.J. [۲۴]  
Florida, Orlando, AIAA, Forum Scitech AIAA ۲۰th in: system," multi-body a in orbit periodic ۲۰۲۰, pp. ۱-۱۹.

meta-learning: reinforcement via guidance adaptive "Terminal Furfaro, R. Linares, R. Gaudet, B. [۲۵]  
Astronaut. Acta operations," close-proximity asteroid autonomous to Applications ۱۷۱ (۲۰۲۰) ۱-۱۳.

unknown with asteroid an over hovering degree-of-freedom "Six Furfaro, R. Linares, R. Gaudet, B. [۲۶]  
Orlando, AIAA, Forum Scitech AIAA ۲۰th in: learning," reinforcement via dynamics environmental ۲۰۲۰, pp. ۱-۱۵, Florida.

[۲۷] صیامی عراقی، م.، شفیعی نژاد، ا. (۱۳۹۹). طراحی مسیر اتصال میان دو ماهواره با الگوریتم یادگیری تقویتی Q. پنجمین  
کنفرانس ملی مهندسی مکانیک و هوافضا. <https://civilica.com/doc/۱۰۳۹۶۳۲>.

[۲۸] کنکاشور، محمدرسول، بلندی، حسین، مزینی، ناصر. (۱۴۰۱). طراحی کنترل‌کننده تطبیقی بهینه بدون مدل برای باز پیکربندی  
آرایش پروازی ماهواره‌ها با یادگیری تقویتی. دانش و فناوری هوافضا، ۱۱(۱)، ۴۱-۶۰.

[۲۹] Farzan Moosavi, (۱۴۰۱). ناوبری و کنترل بازوی رباتیک ماهواره پرواز - آزاد حین جا به جایی و قراردادن محموله در  
ایستگاه فضایی. <http://library.sharif.ir/parvan/resource/۴۹۶۲۹۵>.