



## فرم تعریف پروژه کارشناسی ارشد

تاریخ: ۱۴۰۲/۰۸/۱۰  
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 یادگیری تقویتی مقاوم مبتنی بر بازی دیفرانسیلی در محیط‌های پویای 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. روش‌های هدایت بهینه قدیمی جهت کنترل ماهواره‌ها اغلب نیازمند فرضیات ساده کننده، منابع محاسباتی فراوان و شرایط اولیه مناسب هستند. الگوریتم‌های مبتنی بر یادگیری تقویتی این توانایی را دارند که بدون مشکلات اشاره‌شده هدایت ماهواره را انجام دهند. به همین دلیل، این الگوریتم‌ها می‌توانند امکان محاسبات درونی (On-board Computing) را فراهم می‌کنند.

اهمیت موضوع:

رویکردهای قدیمی برای هدایت فضاپیما در محیط‌های پیچیده چندجسمی اغلب به ایستگاه‌های کنترل زمینی متکی هستند که می‌توانند در معرض خرابی ارتباطات، تاخیرهای زمانی و محدودیت‌های انتقال داده باشند. لذا، ارائه روش‌های هدایت بهینه و مقاوم که امکان محاسبه فرامین هدایت را درون ماهواره فراهم کند، از اهمیت بالایی برخوردار است.

کاربردها:

الگوریتم‌های هدایت مقاوم که محاسبات آن‌ها به صورت درون مداری انجام می‌شود، در ماموریت‌های فضایی مختلفی مانند ماهواره‌ای، سفر به ماه و نزدیک به زمین کاربرد دارند. این الگوریتم‌ها دارای مزایایی مانند استقلال از ایستگاه زمینی، سرعت تصمیم‌گیری و بهبود ایمنی هستند. از کاربردهای این الگوریتم‌ها می‌توان به کنترل مدار و وضعیت ماهواره‌ها، کنترل مدار و وضعیت فضاپیماها در طول سفر به ماه، کنترل مدار و وضعیت فضاپیماها در مدارهای نزدیک به زمین، تعمیر و نگهداری ماموریت‌های 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 سه جسم محدود دایره‌ای Circular Restricted Three-Body Problem (CR3BP) 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



تاریخ: ۱۴۰۲/۰۸/۱۰

شماره: .....

پیوست: .....

الگوریتم هدایت حلقه‌بسته با استفاده از اصول یادگیری تقویتی بر مبنای بازی دیفرانسیلی است که به فضاپیما اجازه می‌دهد با فرض بدترین اغتشاش به مسیر مرجع بازگردد، مسیر مرجع را دنبال کرده و به مدار مقصد برسد. بازی دیفرانسیلی موجب می‌شود الگوریتم هدایت نه تنها بهینه، بلکه مقاوم نیز باشد.

#### فرضیات مسئله:

- مدل دینامیکی مسأله، مسأله سه جسم محدود دایره‌ای سیستم زمین-ماه است. این فرضیه یک سیستم دو جسمی ساده است که فضاپیما را به عنوان جسم سوم بسیار کوچک فرض می‌کند.
- حرکت فضاپیما تنها در صفحه اتفاق می‌افتد و تنها نیروهای گرانشی و اغتشاش در نظر 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.
- محیط یادگیری تقویتی ویژگی‌های مارکوف را دارد؛ به این معنا که حالت فعلی تمام اطلاعات لازم برای پیش‌بینی حالت‌های آینده را فراهم می‌کند.
- فضاپیما در مسیر نزدیک به مرجع با فرض بدترین اغتشاش می‌ماند تا به مدار مقصد برسد.



تاریخ: ۱۴۰۲/۰۸/۱۰

شماره: .....

پیوست: .....

روش انجام کار:

### • مرحله ۱: مدل سازی و شبیه سازی محیط دینامیکی

- پیاده سازی مدل دینامیکی محیط مورد نظر انجام می شود. این مدل باید شامل تمام جوانب مهم محیط از جمله قوانین حرکت، ابعاد فضایی و حالت اولیه باشد.
- مدل دینامیکی محیط به صورتی که اطلاعات مورد نیاز برای یادگیری تقویتی (مانند حالت فعلی) را تامین کند، ایجاد می شود.

### • مرحله ۲: ایجاد و آموزش بازیگر هدایت

- در این مرحله، بازیگری برای انجام هدایت سفینه در محیط ایجاد می شود. این بازیگر مسئول انتخاب اعمال (اعمال کمپیشنر) بر اساس حالت فعلی محیط است.
- از الگوریتم های یادگیری تقویتی مختلف مانند Q-Learning، Deep Q-Networks (DQN)، یا Proximal Policy Optimization (PPO) برای ایجاد بازیگر هدایت استفاده می شود. انتخاب الگوریتم بستگی به محیط و اهداف آموزش دارد.

### پیشینه ی موضوع:

#### پژوهش های خارجی:

هدایت فضاپیماها معمولاً با استفاده از ایستگاه های زمینی انجام می شود. با این حال، این تکنیک ها دارای محدودیت هایی از جمله حساسیت به قطع ارتباطات، تاخیرهای زمانی، و محدودیت های منابع محاسباتی هستند. الگوریتم های یادگیری تقویتی و بازی های دیفرانسیلی می توانند برای بهبود قابلیت های هدایت فضاپیماها، از جمله مقاومت در برابر تغییرات محیطی، کاهش تاخیرهای ناشی از ارتباطات زمینی، و افزایش کارایی محاسباتی، مورد استفاده قرار گیرند.

هدایت فضاپیماها معمولاً پیش از پرواز انجام می شود. این روش ها می توانند از تکنیک های بهینه سازی فراگیر [۱] یا برنامه نویسی غیرخطی برای تولید مسیرها و فرمان های کنترلی بهینه استفاده کنند. با این حال، این روش ها معمولاً حجم محاسباتی زیادی دارند و برای استفاده درون سفینه نامناسب هستند [۲]. یادگیری ماشین می تواند برای بهبود قابلیت های هدایت فضاپیماها استفاده شود. کنترل کننده شبکه عصبی حلقه بسته می تواند برای محاسبه سریع و خودکار تاریخچه کنترل استفاده شود. یادگیری تقویتی نیز می تواند برای یادگیری رفتارهای هدایت بهینه استفاده شود.

شبکه های عصبی ویژگی های جذابی برای فعال سازی هدایت در فضاپیما دارند. به عنوان مثال، شبکه های عصبی می توانند به طور مستقیم از تخمین های وضعیت به دستورهای پیشران کنترلی که با محدودیت های مأموریت سازگار است، برسند. عملکرد هدایت شبکه های عصبی در مطالعاتی مانند فرود بر سیارات [۲۶]، عملیات نزدیکی به سیارات [۲۵] و کنترل فضاپیما با پیشران از دست رفته [۱۳] نشان داده شده است. تازه ترین پیشرفت های تکنیک های یادگیری ماشین در مسائل خودکار سازی درونی به طور گسترده ای مورد مطالعه قرار گرفته اند؛ از پژوهش های اولیه تا توانایی های پیاده سازی. به عنوان مثال، الگوریتم های یادگیری ماشین ابتدایی در فضاپیماهای مریخی نبرد برای کمک به شناسایی ویژگی های زمین شناسی تعبیه شده اند. الگوریتم AEGIS توانایی انتخاب خودکار هدف توسط یک دوربین در داخل فضاپیماهای Spirit، Opportunity و Curiosity را فعال دارد [۵]. در کامپیوتر پرواز اصلی، فرآیند دقت افزایی (Refinement Process) 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. برنامه های آینده برای کاربردهای یادگیری ماشین درون سفینه شامل توانایی های رباتیکی درون سفینه برای فضاپیماهای Perseverance [۸، ۹] و شناسایی عیب برای Europa Clipper [۷] می شود. الگوریتم های یادگیری ماشین پتانسیلی برای سهم مهمی در مأموریت های اتوماسیون آینده دارند.



## فرم تعریف پروژه کارشناسی ارشد

تاریخ: ۱۴۰۲/۰۸/۱۰  
 شماره: .....  
 پیوست: .....

علاوه بر رباتیک سیاره‌ای، پژوهش‌های مختلفی به استفاده از تکنیک‌های مختلف یادگیری ماشین در مسائل نجومی پرداخته‌اند. در طراحی مسیر عملکرد رگرسیون معمولاً مؤثرتر است. به عنوان مثال، از یک شبکه عصبی (NN) در بهینه‌سازی مسیرهای رانشگر کم‌پیشران استفاده شده است [۱۰]. پژوهش‌های جدید شامل شناسایی انتقال‌های هتروکلینیک [۱۱]، اصلاح مسیر رانشگر کم‌پیشران [۱۲] و تجزیه و تحلیل مشکلات ازدست‌رفتن رانشگر [۱۳] می‌شود.

تکنیک‌های یادگیری نظارتی می‌توانند نتایج مطلوبی تولید کنند؛ اما، دارای محدودیت‌های قابل توجهی هستند. یکی از این محدودیت‌ها این است که این رویکردها بر وجود دانش پیش از فرآیند تصمیم‌گیری متکی هستند. این امر مستلزم دقیق‌بودن داده‌های تولیدشده توسط کاربر برای نتایج مطلوب و همچنین وجود تکنیک‌های موجود برای حل مشکل کنونی و تولید داده است.

در سال‌های اخیر، قابلیت 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 هدایت به صورت جلوگیری از شناسایی [۲۱] است. مطالعاتی که فضاپیماهای رانشگر کم‌پیشران را در یک چارچوب دینامیکی چند بدنی با استفاده از RL انجام شده‌است، شامل طراحی انتقال با استفاده از Q-learning [۲۲]، Proximal Policy Optimization [۲۳] و هدایت نزدیکی مدار [۲۴] است.

پژوهش‌های داخلی:

در زمینه هدایت و کنترل فضاپیماها و ماهواره‌ها، پیشرفت‌های زیادی در سال‌های اخیر حاصل شده است. این پیشرفت‌ها عمدتاً ناشی از توسعه روش‌های جدیدی مانند یادگیری تقویتی، کنترل‌های تطبیقی و بینایی کامپیوتر است. این پیشرفت‌ها به بهبود ایمنی و کارایی مأموریت‌های فضایی کمک کرده است. در مرجع [۲۷]، یادگیری تقویتی برای مانور اتصال دو فضاپیما انجام شده است. در مرجع [۲۸]، کنترل‌های تطبیقی برای بازپیکربندی آرایش پروازی ماهواره‌ها پیاده‌سازی شده است. در مرجع [۲۹]، بینایی کامپیوتر برای ناوبری و کنترل بازوی رباتیک ماهواره اجرا شده است.

## اهداف پژوهش:

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



## فرم تعریف پروژه کارشناسی ارشد

تاریخ: ۱۴۰۲/۰۸/۱۰

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 مسئله سه جسم.
	طراحی ساختار محیط سه جسم و بازیکن یادگیری تقویتی و ارتباط بین محیط و بازیکن.
نیمسال اول سال دوم	طراحی مسیر انتقال بهینه بین دو مدار
	شبیه‌سازی دقیق محیط مسأله سه جسم 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
	اصلاح و بهبود عملکرد یادگیری تقویتی در هدایت فضاپیما
تابستان سال دوم	نتیجه‌گیری از تحقیقات و ارائه گزارش نهایی شامل معرفی مسئله، روش‌های استفاده‌شده، نتایج حاصل و پیشنهادات برای تحقیقات آتی

## نتایج مورد انتظار:

- افزایش ایمنی مأموریت به دلیل استفاده از بازی دیفرانسیلی
  - کاهش مصرف سوخت و هزینه به دلیل بهینه‌بودن الگوریتم هدایت
  - هدایت فضاپیما بدون نیاز به ایستگاه زمینی
  - کاهش هزینه ساخت به دلیل عدم نیاز به سیستم‌های ارتباطی قوی
- روش صحت‌سنجی نتایج:
- مقایسه با سایر روش‌های معتبر
  - مدل‌سازی و شبیه‌سازی سیستم محیط و بازیکن
  - گلوگاه‌های پیش‌بینی‌شده:
  - آموزش شبکه: به علت پیچیده‌بودن شبکه و محیطی که طراحی شده است، محاسبات سنگین می‌شود.
  - پایداری و کنترل: به علت غیرخطی بودن دینامیک سامانه و ناپایداری ذاتی سیستم، کنترل و پایداری آن مشکل است.
- نوآوری‌ها:

پیش از این، از ترکیب یادگیری تقویتی و بازی دیفرانسیلی در یک مسئله چندجسمی پویا استفاده نشده است. به علاوه، منطق هدایت ساخته‌شده از ترکیب دو روش، یک هدایت مقاوم و در عین حال بهینه است، که تاکنون در این مسئله انجام نشده است.



تاریخ: ۱۴۰۲/۰۸/۱۰

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 ۲۰۰۴: ۲۱۰-۲۳۳. ۱۰۱۷ Sciences of Academy York New the of Annals System." 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 ۲۰۱۳: (۲۰۱۳) ۸۹ Astronautica Acta manifolds." doi: ۱۵.۰۲۰۲۰۱۳.j.actaastro/۱۰۱۶.۱۰
- [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. ۲۰۱۲, ۱۹-۱, ۳ (TIST), Technol.
- [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 ۲۰۱۹ Alaska, Anchorage, Machinery, Computing ۳۳۳۰۶۵۶.۱۱۴۵/۳۲۹۲۵۰۰.۱۰ doi: , ۲۲۰۱-۲۱۹۱ pp. , ۲۰۱۹
- [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 ۲۰۱۹: 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: ۱۴۰۲/۰۸/۱۰  
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."  
.۱۶-۱ pp. ، ۲۰۰۴ Island.
- networks," neural artificial using connections heteroclinic "Identifying Scheeres. D.J. Smet, De S. [۱۱]  
.۱۹۹-۱۹۲ (۲۰۱۹) ۱۶۱ Astronaut. Acta
- the- (Ph.D. Space Translunar and Cislunar in Optimization Trajectory Thrust "Low Parrish. N.L.O. [۱۲]  
.۲۰۱۸ Boulder, Colorado of University sis),"
- Applica- astrodynamics: in control optimal network "Neural Laipert. F.E. Sood, R. Rubinsztein, A. [۱۳]  
.۲۰۳-۱۹۲ (۲۰۲۰) ۱۷۶ Astronaut. Acta problem," thrust missed the to tion
- 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]  
۱۷۱ 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]  
.۱۷۴۱-۱۷۲۳ (۲۰۲۰) ۶۵ Res. Space Adv. landing,"
- 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,"  
.۱۵-۱ pp. ، ۲۰۱۹ Hawaii, Ka'anapali.
- of guidance intercept angle-only for learning "Reinforcement Linares, R. Furfaro, R. Gaudet, B. [۱۹]  
. (۲۰۲۰) ۹۹ Technol. Sci. Aerosp. targets," maneuvering
- 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  
.۲۰-۱ pp. ، ۲۰۱۹ Maine, Portland, Society, Astronautical
- 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  
.۱۱-۱ pp. ، ۲۰۲۰



تاریخ: ۱۴۰۲/۰۸/۱۰

شماره: .....

پیوست: .....

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

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 (۲۰۲۰) ۱۷۱-۱۳.

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] موسوی، فرزاد (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>