

OpenAI 击败 Dota 2 世界冠军后记：如何训练你的 AI

人工智能2019-04-19



来源：OpenAI

本周末 OpenAI Five 与世界冠军 Dota 2 团队 OG 的总决赛中，它 2 比 0 大败对手。

这意味着 OpenAI Five 是第一个在电竞游戏中击败世界冠军的 AI。

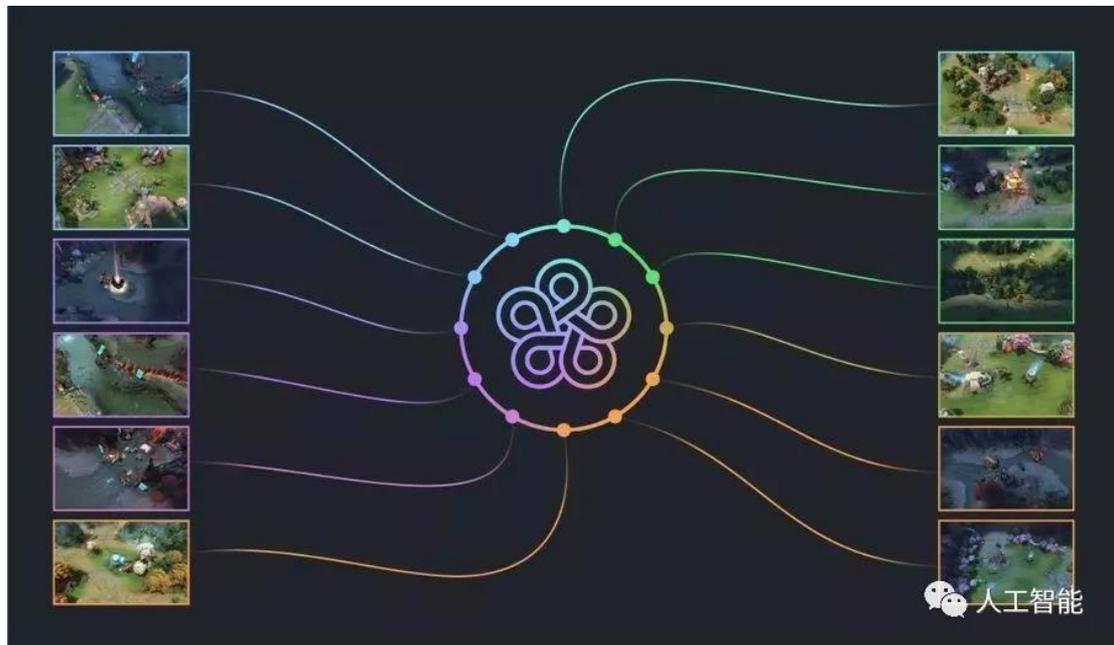
OpenAI Five 和 DeepMind 的 AlphaStar 之前曾私下击败过优秀的职业选手，但却输掉了职业比赛。这次在直播中的表现可以说出乎了大多数人的意料。

在比赛后，OpenAI 也发博文总结了这次比赛的经验以及在也公布了他们在训练 OpenAI Five 过程中的经验教训。一起来看！

在 OpenAI Five 的总决赛中，还有两个惊喜：

1. 尽管训练过程专注于击败其他机器人，但 OpenAI Five 具有了如何成为人类队友的初级能力。这很容易将竞争性 AI 转变为合作型 AI，通过积极的开发工作，未来的 AI 系统会对人类非常有益。

2. 最后的测试将让我们回答一个重要的研究问题：在多大程度上 OpenAI Five 可以被利用或以其他方式被确定地击败。这可能是有史以来最大规模的高强度深度强化学习智能体的部署，人们可以有意识地与之交互。



为什么是 Dota?

启动 OpenAI Five 是为了解决现有深度强化学习算法无法实现的问题。希望通过解决当前方法无法解决的问题，这需要大幅增加我们工具的功能。我们需要复杂的算法思想，例如分层强化学习，但我们对所发现的东西感到惊讶：我们对这个问题所需的基本改进是规模。实现和利用这种规模并不容易，这正是我们研究工作的主要内容！

Human View

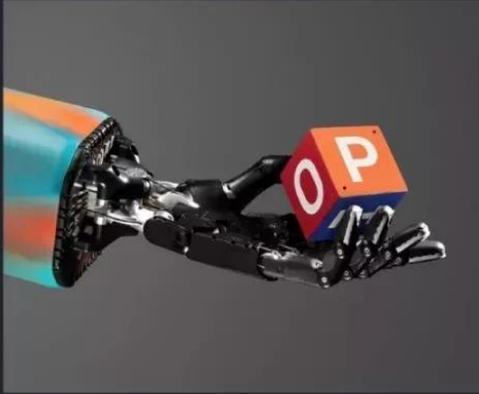


AI View

3.006	-1.386	-0.4695	0.883	1	0.84
-0.3154	-0.5425	-0.5	0.866	0	0.82
3.11	-1.36	-0.9336	0.3584	1	0.78
-2.324	2.863	0.9746	0.225	0	0.86
3.037	-1.361	-0.7773	0.6294	1	0.82
-1.387	2.951	0.988	0.1565	0	0.74
3.023	-0.9395	0.05234	-0.9985	0	0.66
2.951	-0.5747	0.01746	1	0	0.72
2.963	-1.303	0.3906	0.9204	0	0.68
2.834	-3.164	0.01746	-1	0	0.68
3.127	-1.368	0.6562	0.755	1	0.55
3.088	-1.366	0.4695	0.883	0	0.55
2.984	-1.398	-0.225	0.9746	1	0.55
3.037	-1.391	0.788	0.6157	0	0.55
3.076	-1.438	0.883	0.4695	0	0.55
-2.412	2.846	0.996	0.08716	1	0.3

人工智能

Human View



AI View

0.07155341984	0.1856500915	0.192534660
0.06754000613	0.2001686769	0.206590737
0.06228982219	0.199201747	0.21821402
0.05501284074	0.209774121	0.224023983
0.04931758296	0.2110927133	0.226907455
0.04192665512	0.2187263648	0.218435396
0.03458805963	0.2223477091	0.22098078
0.02981736443	0.2164824055	0.223115968
0.02428471393	0.2145174525	0.223008855
0.01733000709	0.2182403815	0.222125609
0.01853848312	0.2234897601	0.22817019
0.02609310462	0.2235220601	0.224757986
0.03343425073	0.2253894907	0.232408747
0.04154509263	0.2246084071	0.230226917
0.04881679852	0.225467511	0.230966722
0.04828479414	0.2271819552	0.228048178

人工智能

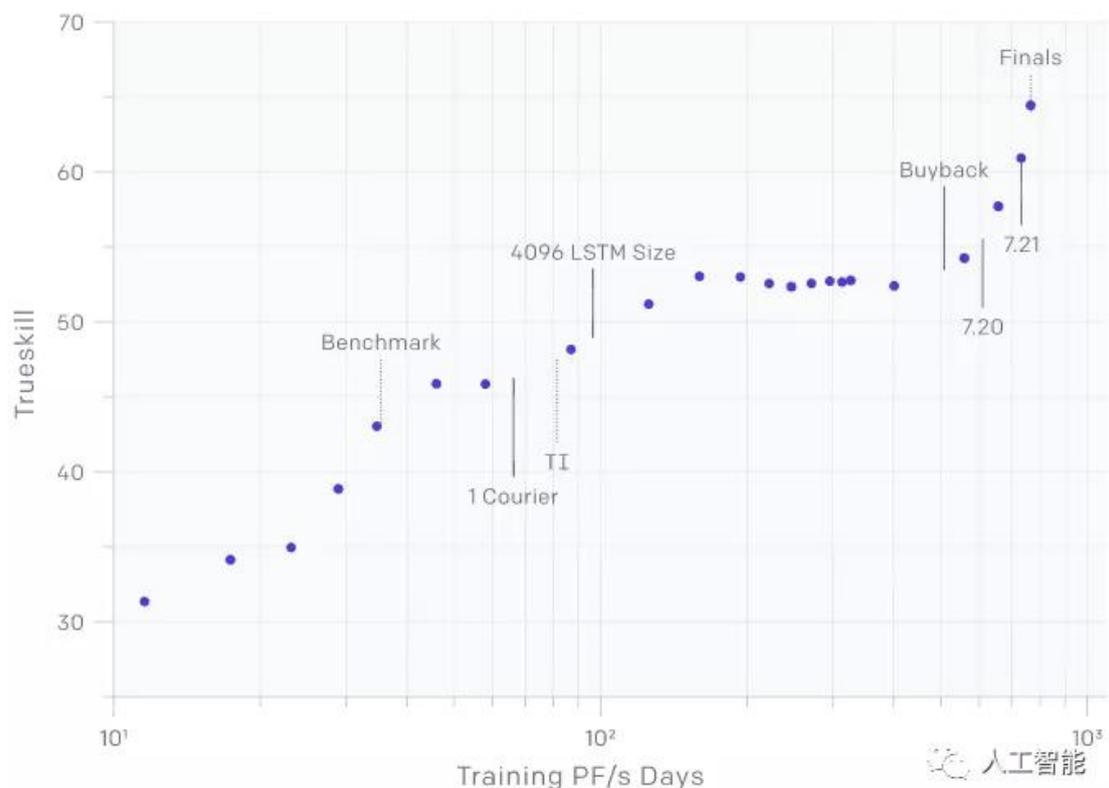
为了构建 OpenAI Five，我们创建了一个名为 **Rapid** 的系统，让我们以前所未有的规模运行 PPO。结果超出了我们最大的期望，我们生产了世界级的 Dota 机器人，没有受到任何基本的性能限制。

当今 RL 算法令人惊讶的强大功能是以大量经验为代价，这在游戏或模拟环境之外是不切实际的。这种限制可能没有声称的那么糟糕，例如，我们使用 **Rapid** 来控制机器人以灵巧地重新定位块，这完全是在模拟中训练并在物理机器人上执行。但我们认为减少经验量是 RL 的下一个挑战。

今天，OpenAI Five 作为竞争者将要退休，但取得的进步和发展的技术将继续驱动我们未来的工作。这不是我们 Dota 工作的结束——我们认为 Dota 对于 RL 开发来说比现在使用的标准环境更具有内在的趣味和困难（现在这已经很好理解！）。

算力

周六 OpenAI Five 的胜利与 2018 年国际赛的失败相比，是由于一次重大变化：训练计算量增加了 8 倍。在项目的许多前期阶段，我们通过提高训练规模来推动进一步的发展。但是在国际赛之后，我们已经将我们项目计算量的绝大部分用于训练单一的 OpenAI Five 模型。所以，我们以唯一可用的方式增加了计算规模：更长久的训练。



OpenAI Five 的 TrueSkill。因为我们已经应用了额外的训练计算，其中线条划分了主要的系统变化（转向单一通道;将 LSTM 大小增加到 4096 单位;升级到补丁版本 7.20 和 7.21）。该图大致是线性的，这意味着 OpenAI Five 从额外的计算中不断受益（注意这是一个双对数图，因为 x 轴是计算量的对数，而 TrueSkill 大致对应于指数进展）。这个图表评估了最终游戏规则（1 个通道，补丁 7.21 等）上的所有机器人——甚至那些在旧游戏规则上训练过的机器人。任何这些之后的陡坡表明 OpenAI Five 适应了这种变化；根据变化，评估可能对之前的版本不公平。

总的来说，当前版本的 OpenAI Five 已经消耗了 800 petaflop / s-days，并且在 10 个实时月内经历了大约 45,000 年的 Dota 自我游戏（从国际赛上的 1.5 个实时月份开始，大约 10,000 年），平均每天 250 年的模拟经验。OpenAI Five 的总决赛版本与 TI 版本相比，胜率为 99.9%。

迁移学习

尽管模型大小和游戏规则发生了变化（包括一些相当大的游戏补丁更新和新实现的功能），但目前版本的 OpenAI Five 自 2018 年 6 月以来一直在不断训练。在每种情况下，我们都能够将模型转移并继续训练，这对于其他领域的 RL 来说是一个开放的挑战。据我们所知，这是 RL 智能体第一次使用如此长期的训练课程进行训练。

为了完成这项工作，我们继续充实我们的诊断工具，以便我们可以从训练有素的参数开始，甚至跨越实质性的架构变化。

更多英雄

我们看到从 5 个到 18 个英雄的训练速度并没有明显减弱。我们假设对更多英雄来说也是如此，在国际赛之后，我们投入了大量精力来整合新的英雄。

我们花了几个星期的时间训练英雄池至 25 个英雄，将这些英雄训练到大约 5k MMR（大约 95% 的 Dota 玩家）。尽管他们还在进步，但他们的学习速度还不够快，无法在总决赛之前达到职业水平。我们尚未有时间调查原因，但我们的怀疑包括模型容量不足，以及需要更好的匹配扩展的英雄池去要求更多的训练时间来让新英雄赶上老英雄。想象一下，当你习惯了某个英雄之后，学习新英雄是多么难！

我们相信这些问题从根本上是可以解决的，解决它们本身就很有趣。在总决赛中的可选英雄中删除了巫妖，可选英雄数量变为 17 位因为巫妖的能力在 Dota7.20 版本中发生了显著变化。

协同模式

它实际上感觉很好；我的冥界亚龙在某些时候为我献出了生命。他试图帮助我，想着“我确定她知道她在做什么”然后显然我没有。但是，你知道，他相信我。和[人类]一起时，我并没有获得这些。

——Sheever



在总决赛期间，我们展示了 **OpenAI Five** 与人类一起在团队中的表现。这场比赛的特色是一队有 **Blitz** 和 **Sheever** 以及由 **Five** 控制的 3 个智能体，面对 **ODPixel**、**Capitalist** 和 3 个由 **Five** 的独立副本控制的智能体。

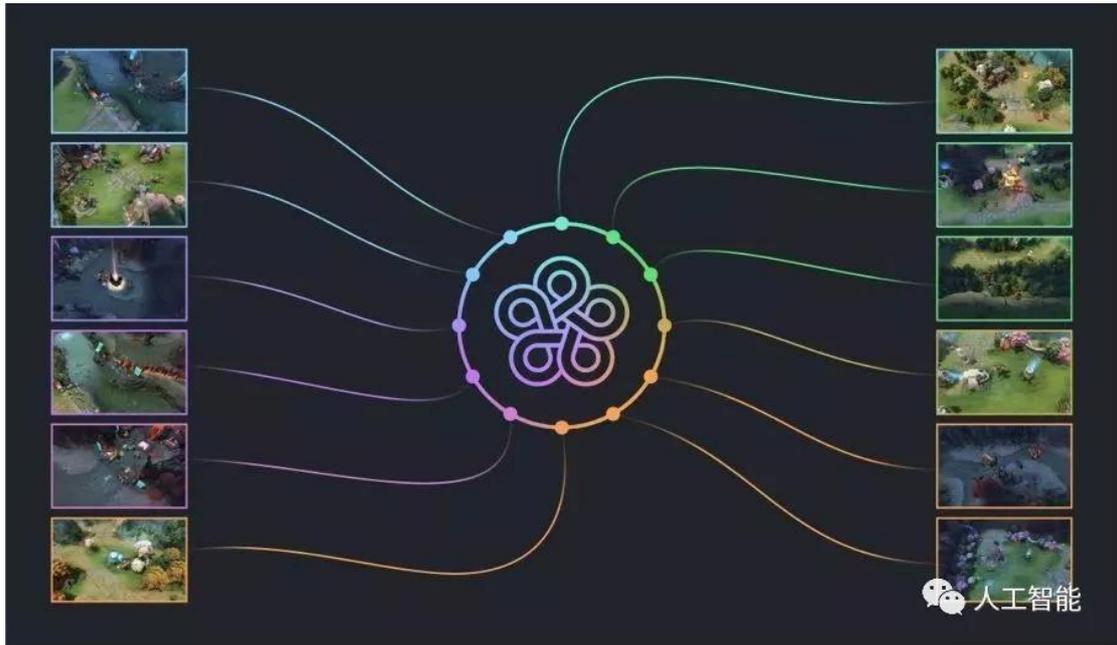
OpenAI Five 与人类玩耍的能力为人类与 **AI** 互动的未来提供了一个很好的愿景，即 **AI** 系统协作并增强人类体验。我们的测试人员报告说，他们感觉得到了机器队友的支持，他们从这些先进的系统中学到了很多，而且整体来说这通常是一种有趣的体验。

请注意，**OpenAI Five** 展示了零样本迁移学习：训练它让所有英雄都受到自身副本的控制，但是可以推广到控制一部分英雄，与人类协作或对抗。我们非常惊讶于这工作得很有效。事实上，我们考虑在国际赛上进行合作比赛，但认为需要专门的训练。

竞技场

我们正在推出 **OpenAI Five** 竞技场，这是一项公共实验，我们将让任何人在竞争和合作模式下玩 **OpenAI Five**。我们知道我们的 1v1 机器人可以通过更聪明的策略来利用；我们不知道 **OpenAI Five** 在多大程度上也是如此，但我们很高兴邀请社区帮助我们找到答案！

竞技场于 4 月 18 日星期四太平洋标准时间下午 6 点开放，并于 4 月 21 日星期日太平洋标准时间晚上 11:59 关闭。



下一步是什么

一旦我们审查了 OpenAI Five 竞技场的结果，我们将发布对 OpenAI Five 的更多技术分析。

之后，我们将继续使用 OpenAI 中的 Dota 2 环境。我们已经看到过去两年 RL 性能取得了快速进展，我们认为 Dota 2 将继续帮助我们推进可行的方案——无论是通过较少的数据还是真正的 AI 合作实现得力的表现。

相关报道：

<https://openai.com/blog/how-to-train-your-openai-five/>

- END -