Phân tích: Đi thẳng theo hướng ngẫu nhiên để tới khu vực mục tiêu nhanh nhất theo kỳ vọng

Dưới đây là một phân tích (kèm chứng minh phác thảo) cho thấy **chọn một hướng ngẫu nhiên rồi đi thẳng** là cách **nhanh nhất theo kỳ vọng quãng đường** để tới đúng "khu vực mục tiêu" trong một thế giới 2D lớn, gồm nhiều khu vực lặp lại (kiểu "biome" như Minecraft), ngay cả khi biên các khu vực rất ngoằn ngoèo.

---

1 Mô hình hoá

- Mặt phẳng được chia thành nhiều **khu vực (biome)**. Gọi T là **khu vực mục tiêu**; phần còn lại là $\overline{T}$.
- Giả thiết **đẳng hướng & đồng nhất** (stationary, isotropic): mô hình không ưu tiên hướng nào, các loại khu vực lặp lại "đều tay"; **tỉ lệ diện tích** (area fraction) của T là $p \in (0, 1)$, của $\overline{T}$ là $1 - p$.
- Gọi $\mathcal{L}_{T,\overline{T}}$ là **mật độ chiều dài biên phân pha** giữa T và $\overline{T}$ (tổng độ dài đường ranh $T|\overline{T}$ trên **một đơn vị diện tích**). Đây là đại lượng chuẩn trong **hình học tích phân / stereology**.

Ta khởi tạo tại một **điểm ngẫu nhiên** trên mặt phẳng.
Mục tiêu: **đến T nhanh nhất** (tối thiểu hoá **quãng đường** đi).

---

2 Ứng viên chiến lược

- **Chiến lược thẳng (Straight-ray):** Chọn $\Theta \sim \text{Uniform}(0, 2\pi)$ rồi **đi thẳng** theo hướng Θ cho đến khi chạm T.
- Các chiến lược "quét" khác như **xoắn ốc**, **zic-zac**, **random walk**... đều có đường đi **dài hơn** để đạt tới **cùng** điểm chạm đầu tiên theo **bất đẳng thức tam giác**.

Mệnh đề tối ưu hoá cục bộ. Với **một hướng θ đã chọn**, đường thẳng từ điểm xuất phát tới **điểm chạm T đầu tiên theo hướng đó** là **đường ngắn nhất**. Mọi quỹ đạo không thẳng đều dài hơn hoặc bằng. $\Rightarrow$ Nếu không có thông tin bản đồ (chỉ "chọn trước" được một hướng), **đi thẳng là tối ưu** trong lớp chiến lược "một hướng, không phản hồi".

---

3 Công thức kỳ vọng: “đi thẳng” tốt tới mức nào?

Đòn bẩy là **công thức dây trung bình (mean chord) trong 2D**. Với một **pha** có tỉ lệ diện tích ϕ trong môi trường đẳng hướng, **độ dài đoạn cắt trung bình** (mean chord) của pha đó bởi **một đường thẳng ngẫu nhiên** là

$$\bar{\ell} = \frac{4\phi}{\mathcal{L}},$$

trong đó $\mathcal{L}$ là **mật độ chiều dài biên phân pha** liên quan đến pha ấy (chiều dài biên trên đơn vị diện tích).

Do khoảng cách theo **một tia** từ điểm ngẫu nhiên **bên trong pha** đến biên đúng bằng **nửa chord** theo tia đó, ta được:

$$\boxed{\mathbb{E}\big[D \,\big|\, X \in \overline{T}\big] = \frac{\bar{\ell}_{\overline{T}}}{2} = \frac{2(1-p)}{\mathcal{L}_{T,\overline{T}}}} \tag{1}$$

Ý nghĩa:

- Nếu ta **chưa ở trong** T (xác suất $1-p$), thì **khoảng cách kỳ vọng** (khi đi thẳng theo hướng ngẫu nhiên) để **chạm ranh** $T|\overline{T}$ – và nhờ đó **vào** T – chính là $\dfrac{2(1-p)}{\mathcal{L}_{T,\overline{T}}}$.
- Nếu ngay từ đầu đã ở T (xác suất p), khoảng cách là 0.

Kiểm thử biên trường hợp kinh điển. Nếu “T” là **một miền lồi hữu hạn** có diện tích A và chu vi P (toàn bộ môi trường chỉ là miền ấy), thì $\mathcal{L}_{T,\overline{T}} = P/A$ và $p = 1$. Công thức (1) suy biến về $\mathbb{E}[D] = 2A/P$ – đúng với định lý Cauchy–Crofton quen thuộc.

---

4 Vì sao thẳng thắng xoắn ốc hay lang thang?

- **Bậc tăng trưởng:** (1) cho thấy “đi thẳng” có kỳ vọng **tuyến tính** theo **tỉ số hình học** $(1-p)/\mathcal{L}_{T,\overline{T}}$.
- **Xoắn ốc Archimedes** $r = k\theta$: để vươn tới bán kính R cần cung dài xấp xỉ $\sim R^2/(2k)$ (**bậc hai** theo cự ly).
- **Random walk** (bước ngắn, đổi hướng liên tục): quãng đường để đạt dịch chuyển thẳng hàng R có chi phí xấp xỉ **bậc hai** do hiện tượng **lan toả** (diffusion-like).

$\Rightarrow$ Về **kỳ vọng quãng đường**, mọi chiến lược “quét”/“lang thang” đều **thua bậc** so với “đi thẳng”.

---

5 Trường hợp nhiều loại khu vực (ví dụ 5 loại, tỉ lệ bằng nhau)

Khi có nhiều pha $T_1, \dots, T_m$ với $\sum_i p_i = 1$ và muốn tới một pha đích T có tỉ lệ p, ta thay $\mathcal{L}_{T,\overline{T}}$ bằng tổng chiều dài biên **giữa T và mọi pha còn lại** trên đơn vị diện tích. Công thức (1) **giữ nguyên**:

$$\mathbb{E}\big[D \,\big|\, X \in \overline{T}\big] \;=\; \frac{2(1-p)}{\mathcal{L}_{T,\overline{T}}}.$$

Với **5 loại bằng nhau** $\Rightarrow p = \frac{1}{5}$. Khi **mạng ranh giới dày đặc** (nhiều rìa T kề $\overline{T}$), $\mathcal{L}_{T,\overline{T}}$ **lớn** $\Rightarrow$ kỳ vọng **ngắn**; còn nếu T xuất hiện thành "mảng to, cách xa nhau" thì $\mathcal{L}_{T,\overline{T}}$ **nhỏ** $\Rightarrow$ kỳ vọng **dài**. Điều này phản ánh đúng trực giác "đường biên càng dày, càng dễ chạm".

Ví dụ "lưới ô vuông" cạnh L với pha mục tiêu phân bổ ngẫu nhiên độc lập theo tỉ lệ p: có thể ước tính $\mathcal{L}_{T,\overline{T}}$ tỉ lệ $\frac{1}{L}$ nhân với xác suất hai ô kề khác pha $\sim \frac{2p(1-p)}{L}$. Khi đó

$$\mathbb{E}\big[D \,\big|\, X \in \overline{T}\big] \approx \frac{2(1-p)}{2p(1-p)/L} = \frac{L}{p},$$

tuyến tính theo kích cỡ "mảng" (chỉ là xấp xỉ, nhưng đúng về **bậc**).

---

6 Kết luận & câu trả lời "nên làm gì"

Chiến lược tối ưu (không có bản đồ, không tín hiệu dẫn đường): Chọn ngẫu nhiên một hướng và đi thẳng.

- Đây là **tối ưu cục bộ** cho mọi hướng đã chọn (do bất đẳng thức tam giác), và **tối ưu về bậc** so với các chiến lược quét/khảo sát.
- **Kỳ vọng quãng đường (khi chưa ở trong T)** được cho bởi công thức **chặt** (trong khung đẳng hướng, đồng nhất):

$$\boxed{\mathbb{E}\big[D \,\big|\, X \in \overline{T}\big] \;=\; \frac{2(1-p)}{\mathcal{L}_{T,\overline{T}}}}$$

  trong đó $\mathcal{L}_{T,\overline{T}}$ là **mật độ chiều dài ranh $T|\overline{T}$** (đơn vị: chiều dài trên đơn vị diện tích).

Tóm lại: trong một thế giới 2D nhiều khu vực lặp lại, **đi thẳng theo một hướng ngẫu nhiên** là cách **nhanh nhất theo kỳ vọng** để vào khu vực mong muốn. Công thức (1) cho ta thước đo định lượng rõ ràng – phụ thuộc **chỉ** vào tỉ lệ diện tích pha đích p và **mật độ ranh $\mathcal{L}_{T,\overline{T}}$**.