Single Model BUT Multiple Mechanisms

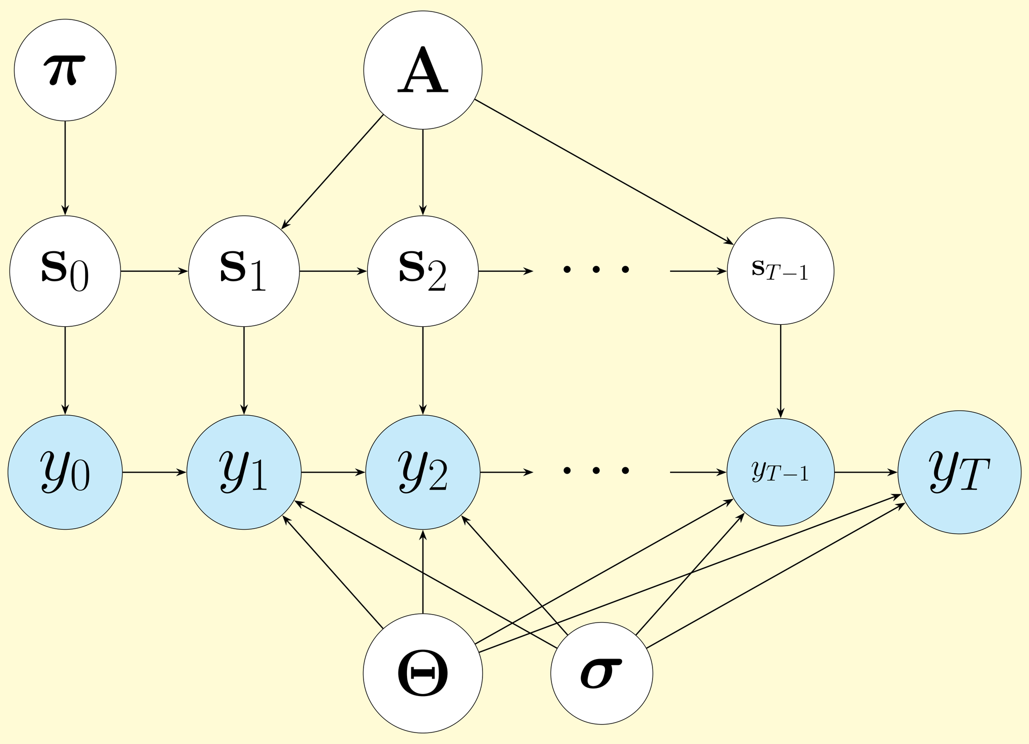
Hybrid Modeling of Hidden Markov Model and Ordinary Differential Equations

Yuhei Yamaguchi (D1, Graduate School of Arts and Letters, Tohoku University) and Yuri Ono (D1, Graduate School of Sociology, Kwansei Gakuin University)

概要・目的

社会現象は単一の方程式で表現できるものであるとは限らない．観測値（データ）が生成されるメカニズムは，時間とともに変化する場合がある．本研究では，離散化した微分方程式と隠れマルコフモデルを統合することで，時変的なデータ生成メカニズムを柔軟に表現する手法と実用例を紹介する．

モデルの構造



- 微分方程式の離散化（とりあえずオイラー法）

$$\frac{dy}{dt} = f(\boldsymbol{\Theta}) \simeq \frac{y_{t+\Delta t} - y_t}{\Delta t}$$
$$\Rightarrow \hat{y}_{t+\Delta t} \simeq y_t + \Delta t \cdot f(\boldsymbol{\Theta})$$

- データ$y_{t+\Delta t}$の尤度

$$p(y_{t+\Delta t}|y_t, \boldsymbol{\Theta}, \sigma) = \mathcal{N}(y_{t+\Delta t}|\hat{y}_{t+\Delta t}, \sigma)$$

- 潜在変数$\mathbf{s}$はK種類ある．最初の潜在変数s_0は$\boldsymbol{\pi}$によって決まる．それ以降の潜在変数は遷移確率行列$\boldsymbol{A}$で決まる．
- τ時点の潜在変数$s_\tau = k$の場合，$\hat{y}_{\tau+\Delta t}$は$\boldsymbol{\Theta}_k$に基づいて計算される．

$$\hat{y}_{\tau+\Delta t} \simeq y_\tau + \Delta t \cdot f_k(\boldsymbol{\Theta}_k)$$

$y_{\tau+\Delta t}$の尤度はσ_kを用いて，

$$p(y_{\tau+\Delta t}|y_\tau, \boldsymbol{\Theta}_k, \sigma_k) = \mathcal{N}(y_{\tau+\Delta t}|\hat{y}_{\tau+\Delta t}, \sigma_k)$$

パラメタ推定

- 同時分布

$$p(\mathbf{y}, \mathbf{s}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{\Theta}, \boldsymbol{\sigma}) = p(\boldsymbol{\pi})p(\mathbf{A})p(\boldsymbol{\Theta})p(\boldsymbol{\sigma})p(y_1|y_0, s_0, \boldsymbol{\Theta}, \boldsymbol{\sigma})p(s_0|\boldsymbol{\pi}) \times \prod_{t=2}^{T} p(y_t|y_{t-1}, s_{t-1}, \boldsymbol{\Theta}, \boldsymbol{\sigma})p(s_{t-1}|s_{t-2}, \mathbf{A}).$$

- パラメタの事後分布

$$p(\boldsymbol{\pi}, \mathbf{A}, \boldsymbol{\Theta}, \boldsymbol{\sigma}|\mathbf{y}) = \frac{p(\mathbf{y}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{\Theta}, \boldsymbol{\sigma})}{p(\mathbf{y})} = \frac{\sum_{\mathbf{s}} p(\mathbf{y}, \mathbf{s}, \boldsymbol{\pi}, \mathbf{A}, \boldsymbol{\Theta}, \boldsymbol{\sigma})}{p(\mathbf{y})}$$
$$\propto \underbrace{p(\boldsymbol{\pi})p(\mathbf{A})p(\boldsymbol{\Theta})p(\boldsymbol{\sigma})}_{\text{prior distributions}} \times \underbrace{\sum_{\mathbf{s}} \left[p(y_1|y_0, \mathbf{s}_0, \boldsymbol{\Theta}, \boldsymbol{\sigma})p(\mathbf{s}_0|\boldsymbol{\pi}) \prod_{t=2}^{T} p(y_t|y_{t-1}, \mathbf{s}_{t-1}, \boldsymbol{\Theta}, \boldsymbol{\sigma})p(\mathbf{s}_{t-1}|\mathbf{s}_{t-2}, \mathbf{A}) \right]}_{\text{Likelihood } p(\mathbf{y}|\boldsymbol{\pi}, \mathbf{A}, \boldsymbol{\Theta}, \boldsymbol{\sigma})}.$$

同時分布を潜在変数について周辺化すると，パラメタの事後分布は**事前分布と尤度の積**で表現できる．

- MCMCでパラメタ推定（**cmdstanr** ver. 0.8.0, **CmdStan** ver. 2.36.0）.

実社会データへの応用例① ファンヒーターの普及率

■定数変動モデル

普及率（y_t）の変動を定数としたモデル．方程式は，$\dfrac{dy_t}{dt} = \lambda_k$.

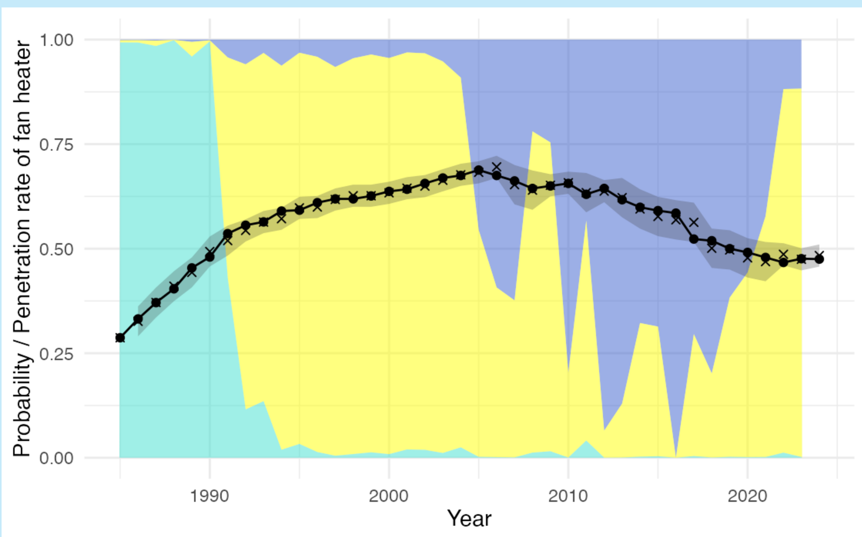


Figure: 3つの潜在変数を仮定した場合

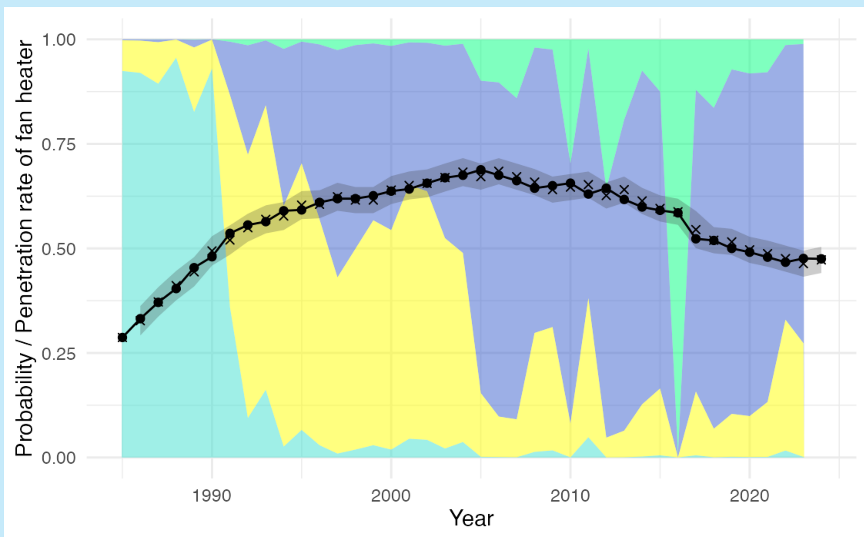


Figure: 4つの潜在変数を仮定した場合

潜在変数3つのときは，強上昇期，弱上昇期，下降期に分離できている．潜在変数を4つに増やすと，2016年から2017年にかけての急下降も検出．

■ 灯油価格指数を考慮したロジスティックモデル

灯油価格指数（p_t）が，ファンヒーター所有の効用を決める．

$$\frac{dy_t}{dt} = y_t(1 - y_t)\lambda_{tk}, \ \ \lambda_{tk} = \beta_{k0} + \beta_{k1}p_t$$

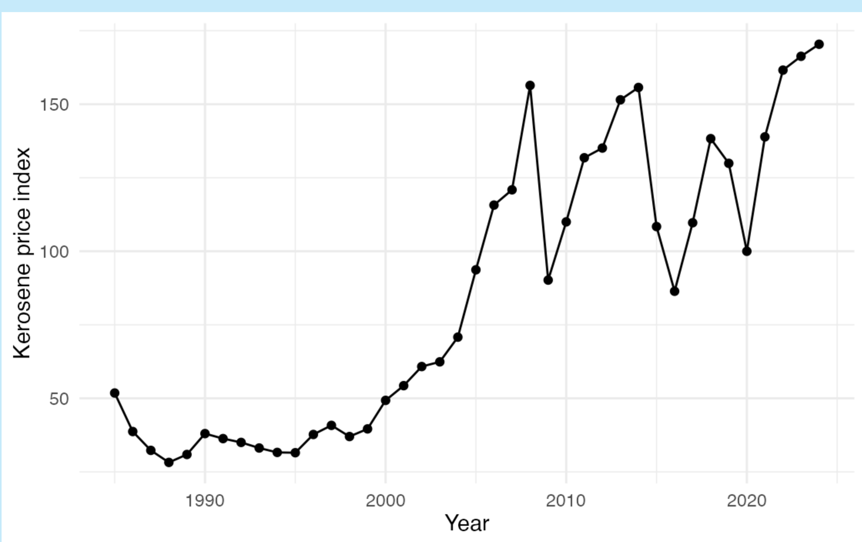


Figure: 灯油価格指数（source: GD Freek!）

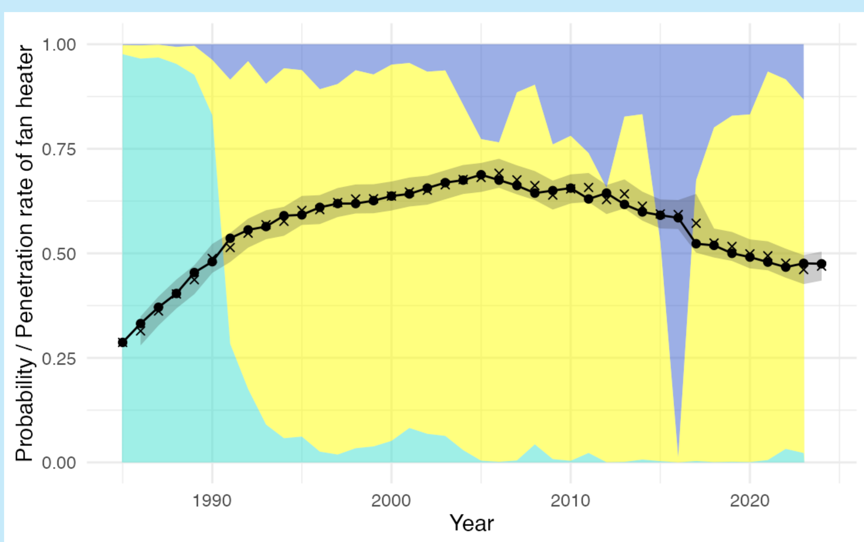


Figure: 3つの潜在変数を仮定した場合

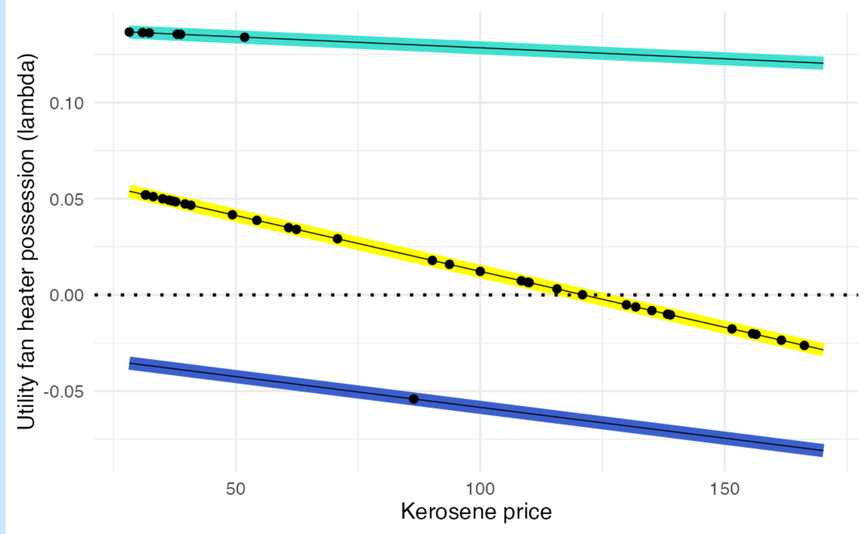


Figure: 潜在変数別，ヒーター購入の効用

- すべての潜在状態で，灯油価格指数が上がるほど，ファンヒーターを所有することの効用（λ）は小さくなる（右下図）
- 1991年以降は，灯油価格が上がると効用が大きく減少しやすい．黄色の潜在状態において，灯油価格が上がると，ヒーター所有の効用は負に転じる．

実社会データへの応用例② 生活保護受給率の推移

- 生活保護を受給せずに働く場合の効用の期待値$\bar{u}_t$の定義

$$\bar{u}_t = \begin{cases} \sum_{m=0}^{N}\{\underline{(\lambda_1 + 1)^m - 1}\}Bin(m; N, q_t) & \text{if } k = 1, \text{(指数型効用関数)} \\ \sum_{m=0}^{N}\{\underline{\lambda_2 \cdot \ln(m+1)}\}Bin(m; N, q_t) & \text{if } k = 2, \text{(対数型効用関数)} \\ \sum_{m=0}^{N}\{\underline{\lambda_3 \cdot m}\}Bin(m; N, q_t) & \text{if } k = 3. \text{(線形効用関数)} \end{cases}$$

ただし，q_tは求職活動に成功する確率で，p_tをt時点の完全失業率として，$q_t = 1 - p_t^{1/N}$．Nは求職に挑戦する回数．mは成功回数

- 生活保護を受給する効用をv，受給者割合をy_tとして，微分方程式は，

$$\frac{dy_t}{dt} = y_t\{v - (y_t v + (1 - y_t)\bar{u}_t)\}.$$

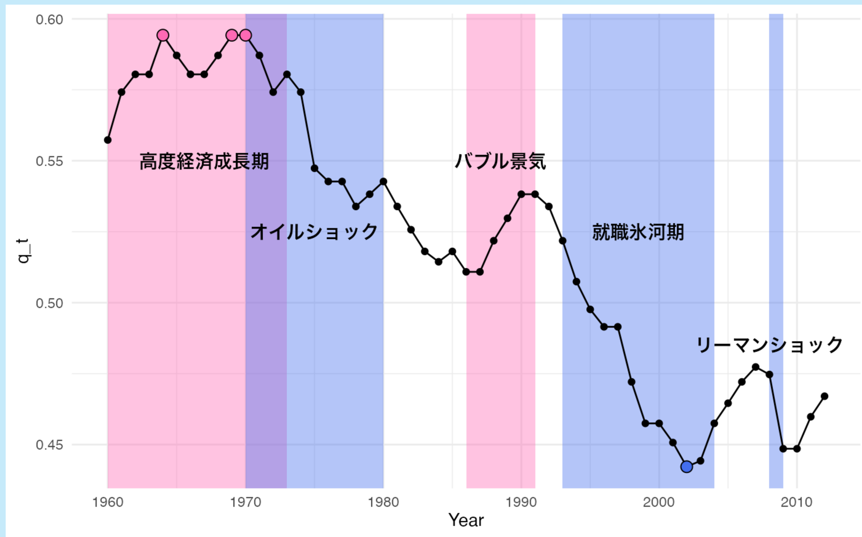


Figure: 求職成功確率q_tの推移

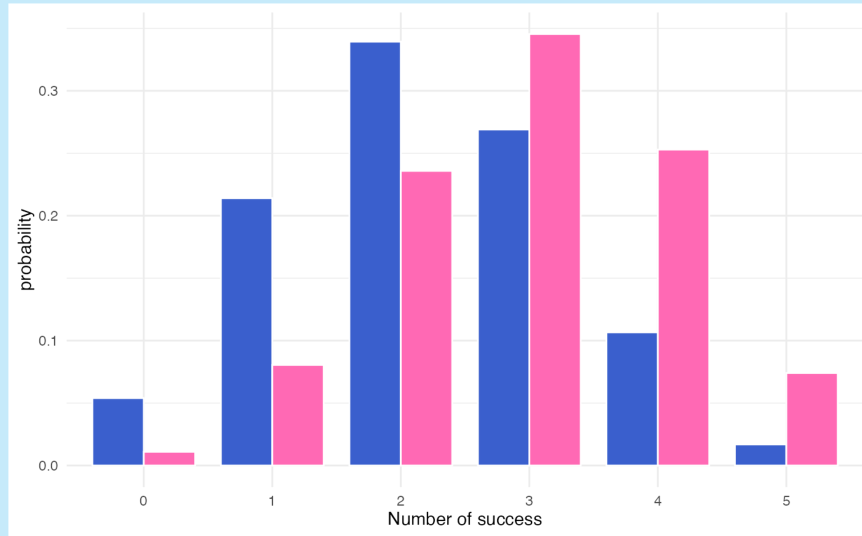


Figure: 求職成功回数の分布例($q = .44$と.59)

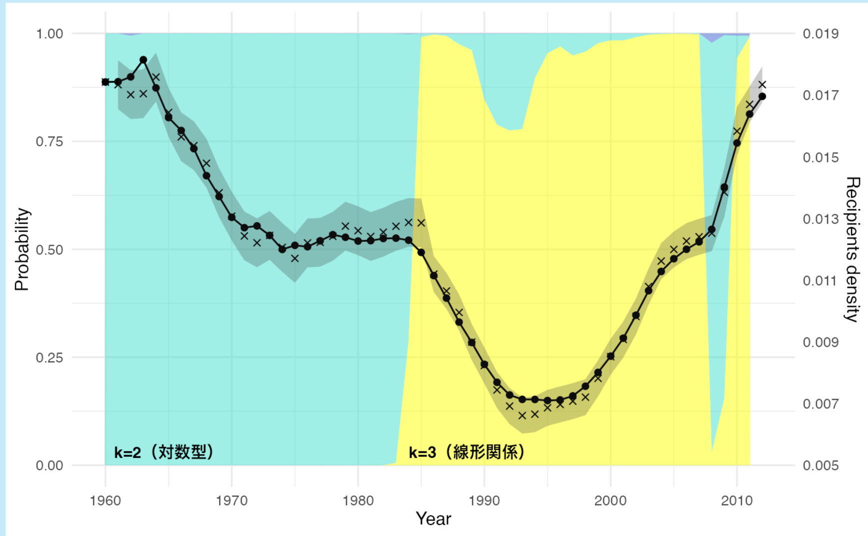


Figure: 結果（$v = 1, N = 5$）

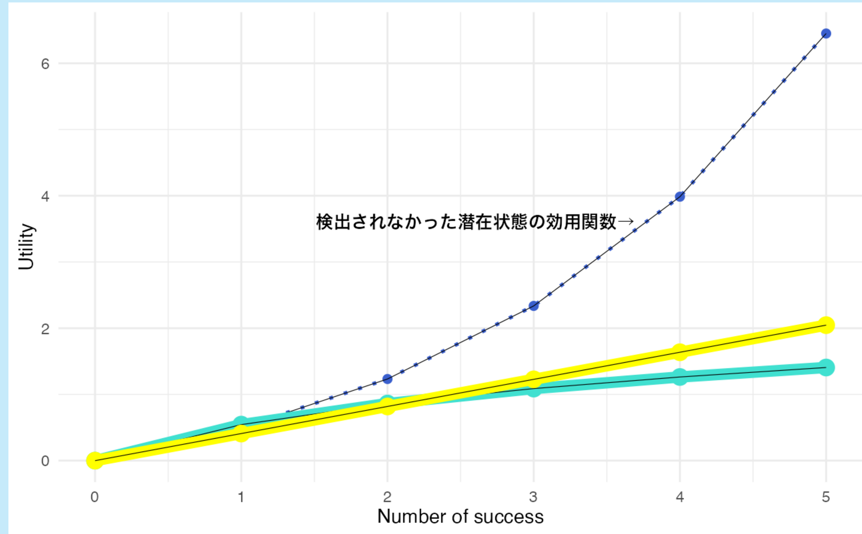


Figure: 潜在変数別，成功回数と効用の関係

- **2つの潜在変数が検出された**
  - ⇒ 生活保護受給率の背後には異なる2つの効用関数がある 求職成功確率が高い時期と低い時期で異なる効用構造が検出された
- **水色の状態（$k = 2$）：対数型効用関数（限界効用逓減）**
  - ⇒ たくさん成功しても，効用は上がりにくい 求職成功回数により生じる効用の格差が小さい社会
- **黄色の状態（$k = 3$）：線形効用関数**
  - ⇒ 求職に成功するほど，どんどん効用が増える 求職成功回数による効用の格差が生じやすい社会
- 1960年–1970年および1985年–1990年は，ともに生活保護受給率が低下
  - ⇒ その背後にあるメカニズム（社会の効用構造）は異なる
- リーマンショック（2008）時には，$k = 2$が採択
  - ⇒ たくさん求職に成功しても効用が上がりにくい社会が再来

まとめ

- **隠れマルコフモデルと微分方程式を統合することで，人間が直接は観測できないデータ生成メカニズムの切り替わりを検出**できる！
- 同じ社会現象であっても，その背後にあるメカニズムは異なっている場合がある．（うまくやれば）本モデルはそれを検出できる！
- **複数ある微分方程式の関数形が異なっていても良い！**
  - ⇒ 連続潜在変数を仮定する状態空間モデルとの差別化（最大の強み）
- **Tips**: MCMCがうまく収束しない場合は，強い事前分布を置いたり，パラメタの大小関係に制約をつけるといい
- **Idea**: 教育と達成地位の連関の強さ（ED連関）の変遷を捉える研究
- **Idea**: 複数のフェイズがある社会現象（e.g. 流行現象）の，各フェイズ分離する研究

Enjoy!