Learning Human Preferences

Owain Evans
Natalie Deschamps shared VR-Zone's photo.

Skateboard baby stroller that comes with brakes and handlebars for steering.

Like VR-Zone for more amazing stuff.

Angela Azarai likes this.

Angela Azaria Hilarious.
2 minutes ago · Like

Write a comment...

Christine Tripp at The Centurion Conference & Event Center
Looking forward to seeing Cara!

Like · Comment · Share · 11 minutes ago in Ottawa ·

Write a comment...

Sara Karissa
Facebook

• 1 billion users per day

• 4.5 billion ‘Likes’ per day

• 10 billion messages per day
Facebook Newsfeed

• Surface: dashboard with updates from friends.

• Underneath: automated system for individualized preference inference.

• Goal: Take 1500 posts from friends, rank them based on which you will like and engage with.

• **Preference inference:** from friends, likes, links followed (+ collaborative filter), *infer* what you’ll like.
Thesis topic

- FB newsfeed: infer individual preferences from FB behavior (and ‘testimony’).

- Thesis: Formal methods of inferring preferences from observed choices (not testimony).

- Goals: predict long-term behavior, understand content of preferences, provide advice/recommendations.

This talk:
(1) Can we learn preferences from choices?

(2) Preferences and well-being. What are the limitations of providing advice using such methods?
1. Can we learn preferences from choices?

Background to this work:

• Decision theory: choices over lotteries -> utility function on money (or apples vs. oranges).

• Economics: ‘revealed preference’

• Machine Learning: Inverse Reinforcement Learning (i.e. ‘utility function inference’)

1. Can we learn preferences from choices?

What is learned from choices?

*Everyday vs. economic* concept of preference.

“Fred is at dinner. He’d *prefer* to drink wine. But he’s promised a friend he’d drive him home. So Fred *chooses* to drink water.” (Hausman 2012)

**Choice**- need to weigh up all dimensions of evaluation: immediate pleasure, moral constraints, commitments (e.g. promises).

Preferences as ‘total comparative evaluations’.
1. Can we learn preferences from choices?

Economic preferences:

- total comparative evaluations

- defined on possible worlds (cf. ‘I prefer the color blue’).

- Model of choice: consider all future consequences of each action, evaluate consequences across all dimensions, pick the best action.
1. Can we learn preferences from choices?

Problem: people don’t always act on their (informed) preferences.

Distinguish:

1. **Economic preference**: all-things-considered judgment about which possible worlds are better (independent of beliefs, plans, time).

2. **Cause of choice**: all-things-considered judgment about best choice. Depends on preferences, plans to realize preferences, beliefs about world, biases/inconsistencies, etc.
1. Can we learn preferences from choices?

Example:
“John is choosing a place to eat. There’s a vegetarian cafe and a donut shop nearby. He ends up eating at the donut shop.”

Inferences:
• He prefers donuts.

• He thought the cafe was closed (or didn’t know about it). **FALSE BELIEF**

• He intended to go the cafe, but when he walked right by the donut shop, he couldn’t resist going in. **TIME INCONSISTENCY**
1. Can we learn preferences from choices?

Problem: if we don’t know about biases/false beliefs, can’t always infer preferences.

Idea: Explicitly model sequential planning with inaccurate beliefs and time inconsistency (hyperbolic discounting). Infer all parameters (preferences + beliefs).

Key observation: False beliefs and time inconsistency produce distinctive behaviors and can be inferred (via Bayesian inference) from choices alone.
1. Can we learn preferences from choices?

Example 1: Walk in direction of closed restaurant (false belief).

Example 2: Remove one cookie from jar. Put jar away. Then repeat. (Time inconsistency)

Example 3: Odysseus tying himself to mast. (Predicted time inconsistency)
2. Preference learning and welfare

You learn preferences and advise choices that maximize them. Is this good for person?

• (informed) preference satisfaction theory of well-being (vs. objective list view)

• Risk: easier to infer ‘bad’ preferences. (Avoid using ideas just discussed).
2. Preference learning and welfare

What if something is good but person doesn’t prefer it (i.e. value it)?

• You can’t infer from their choices. (Maybe from someone else’s choices.)

• Recommend it: might work (based on good reputation of recommender) but will be difficult to influence behavior.