Goal Recognition in Latent Space

Leonardo Amado∗, João Paulo Aires∗, Ramon Fraga Pereira∗
Mauricio C. Magnaguagno∗, Roger Granada∗ and Felipe Meneguzzi†
Pontifical Catholic University of Rio Grande do Sul (PUCRS) – Brazil
✉ {leonardo.amado, joao.aires.001, ramon.pereira, mauricio.magnaguagno, roger.granada}@acad.pucrs.br
☎ felipe.meneguzzi@pucrs.br

Motivation and Goals

• Goal recognition approaches have progressively relaxed requirements regarding:
  – amount and accuracy of domain knowledge; and
  – amount and accuracy of available observations at recognition time.

• However, to recognize goals using raw data, to infer possible goals recent approaches need either:
  – near-flawless human engineered domain knowledge
  – samples of behavior that account for almost all actions being observed

• This is too strong for most real-world applications.

• We develop an approach that leverages advances in recurrent neural networks to perform goal recognition as a classification task.

Main contributions.

• An end-to-end machine learning technique for goal recognition based on training an LSTM network.

• Comparison of the performance and trade-offs of resulting approach with traditional goal recognition approaches.

Goal Recognition in Latent Space

Goal recognition in Latent Space is a technique to apply classical goal recognition algorithms in raw data (such as images) by converting it into latent space. It consists of four key processes:

1. Train an autoencoder capable of creating a latent representation to a state of such image domain.

2. Derive a PDDL domain, by extracting the transitions of such domain when encoded in latent space, obtaining a domain D.

3. Convert to a latent representation a set of images representing the initial state I, the set of facts F, the observations O, and a set of possible goals G, where the hidden goal G̃ is included.

4. Apply goal recognition techniques using the computed tuple ⟨D, F, I, G̃, O⟩.

Our goal is to obviate the need of the second process by avoiding the intermediary PDDL representation and train an LSTM to recognize goals, instead of using classical goal recognition algorithms.

LSTM Approach

At the center of our approach, we use an LSTM network to infer goals that reasons over an auto-encoded representation of the domain:

![LSTM Architecture](image)

This LSTM lies at the center of a latent goal recognition architecture, obviating the need of a PDDL domain.

The inferred goal, in the encoded representation, can be decoded to an image representation.

Experiments

To validate our approach, we use six domains from three distinct games. We compare our approach to the state-of-the-art in goal recognition in latent space.

![Figure 1: LSTM architecture](image)

The inferred goal, in the encoded representation, can be decoded to an image representation.

![Figure 2: Latent goal recognition structure](image)

The throughput, in the encoded representation, can be decoded to an image representation.