Leela AI:
Creating resilient intelligences

Steve Kommrusch, PhD, MIT MSc '89, Senior AI Scientist
Henry Minsky, MIT MSc '93, CTO
Leela AI, Inc.

Two key problems in AI today

Three challenges for Deep Learning
- Deep Supervised Learning works well for perception
  - When labeled data is abundant.
- Deep Reinforcement Learning works well for action generation
  - When trials are cheap, e.g. in simulation.
- **Three problems the community is working on:**
1. **Learning with fewer labeled samples and/or fewer trials**
   - Self-supervised learning / unsup learning / learning to fill in the blanks
     - learning to represent the world before learning tasks
2. **Learning to reason**, beyond "system 1" feed-forward computation.
   - Making reasoning compatible with gradient-based learning.
3. **Learning to plan complex action sequences**
   - Learning hierarchical representations of action plans

- At AAAI in 2020, Yann LeCun highlighted the importance of:
  1) self-supervised learning
  2) complex action planning

We address both problems with an approach inspired by Jean Piaget

Piaget proposed a powerful model of cognitive development that emphasized the hierarchical and sequential development of intelligence.

- Sensorimotor stage: 0-2 years
  - Simple reflexes (moving hands, eyes, etc)
  - Primary circular reactions: coordination of 2 types of schema: i.e. passing hand before face
  - Secondary circular reactions: actions involving external objects begin
  - Coordination of 2nd circular reactions: 'first proper intelligence'; means and ends; goals; object permanence
  - Tertiary circular reactions: curiosity about object properties
  - Internalization of schemas: insight, creativity, use primitive symbols
- Pre-operational stage: 2-7 years
  - Child can form stable concepts and magical beliefs; increased play
- Concrete operational stage: 7-11 years
  - Child can think logically, understand reversibility, see viewpoints of others
- Formal operational stage: 11-16+ years
  - Development of abstract reasoning; utilize metacognition; multistep problem solving
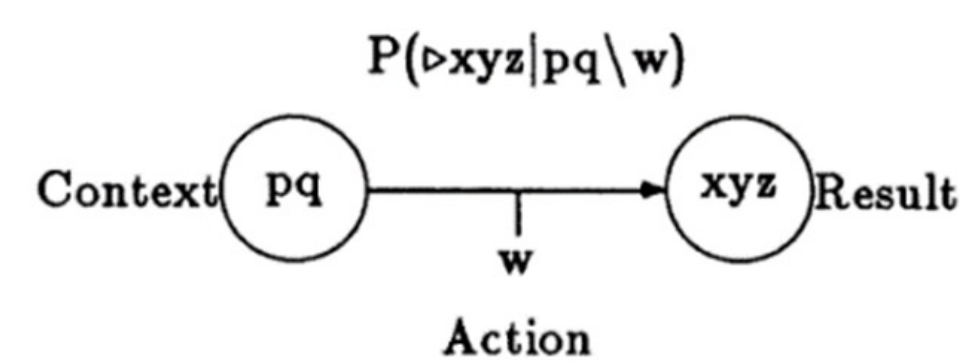
Leela implementation of Piaget's schemas

$$P(\triangleright xyz|pq\backslash w)$$

Context $\bigcirc pq \xrightarrow{\quad w \quad} \bigcirc xyz$ Result

Action

Figure 1: A schema.

Figure 1 shows a schema with items p and q in its context, action w, and items x, y, z in its result. A schema also maintains some auxiliary data, such as the schema's reliability— that is, the reliability with which the predicted result will actually follow the schema's action (provided that the context is satisfied). Reliability is measured by recording:

- $P(\triangleright R|CA)$, the conditional probability of a transition to the result state (R) given context conditions (C) and action (A).
- $P(\triangleright R|C\neg A)$, the conditional probability of a transition to the result under the same conditions except without the action.

- The syntax P($xyz|pq\backslash w$) indicates the conditional probabilities (frequencies) of transitioning to the result of $x,y,$ and z being true, given the condition that p and q are true and action w was taken.

Problem: Time-To-Value for AI Video Analytics is Currently Measured in Months
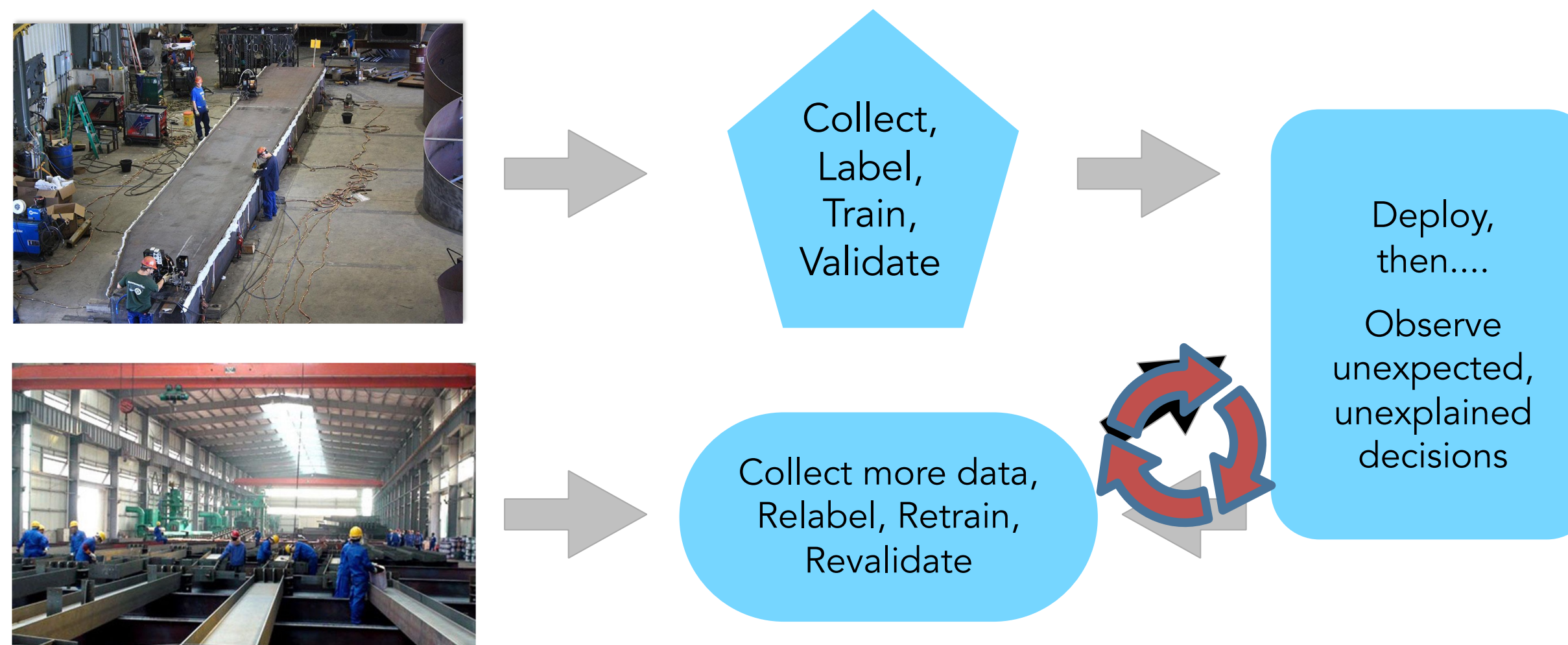
This cycle can take months before any value is delivered!

Understanding video: A No-Code Human-Teachable Solution

Our product shortens the time-to-value cycle to days

- Training requires 100x less data
- Custom use-cases built in hours
- Solutions are robust across locations and cameras
- Users can easily adjust/improve the system
- Decisions are explainable

Benefits seen at customer site

- Understanding where and when non-value-add activities are taking place
- Comparing efficiency of alternate workflows for manufacturing
- Digital Transformation/Industry 4.0 promises kept: Better resource management for labor-intensive industries

Video annotation of objects, actions, and people

- Leela creates explainable summaries of activity over time

Future research opportunities

- Create 3D model based on video
  - Allow simulated actor to learn about results of actions
  - Use action understanding to control robotic behaviors on a dynamic manufacturing floor with complex tasks.

- Improve manufacturing processes
  - Automatically generate suggestions for ways to improve processes, automatic anomaly detection
  - Use understanding of goals and action results to propose improvements
  - Recognize which actions increase risk of injury and detect them in video or adjust process to avoid risk

- Leverage constructivist AI concepts and approach
  - Generalize new concepts by observing action effects on world
  - Use self-supervised methods to improve machine learning
  - Provide more high-level explainable results for human-computer interaction

References

C. Shaoul Leela AI CEO speaks at MIT Manufacturing Conference
https://leela.ai/2022/04/04/leela-ai-ceo-speaks-at-mit-manufacturing-conference/

S. Kommrusch, H. Minsky, M. Minsky, C. Shaoul. Self-Supervised Learning for Multi-Goal Grid World: Comparing Leela and Deep Q Network Proceedings of Machine Learning Research 131: 81-97, 2020
http://proceedings.mlr.press/v131/kommrusch20a/kommrusch20a.pdf