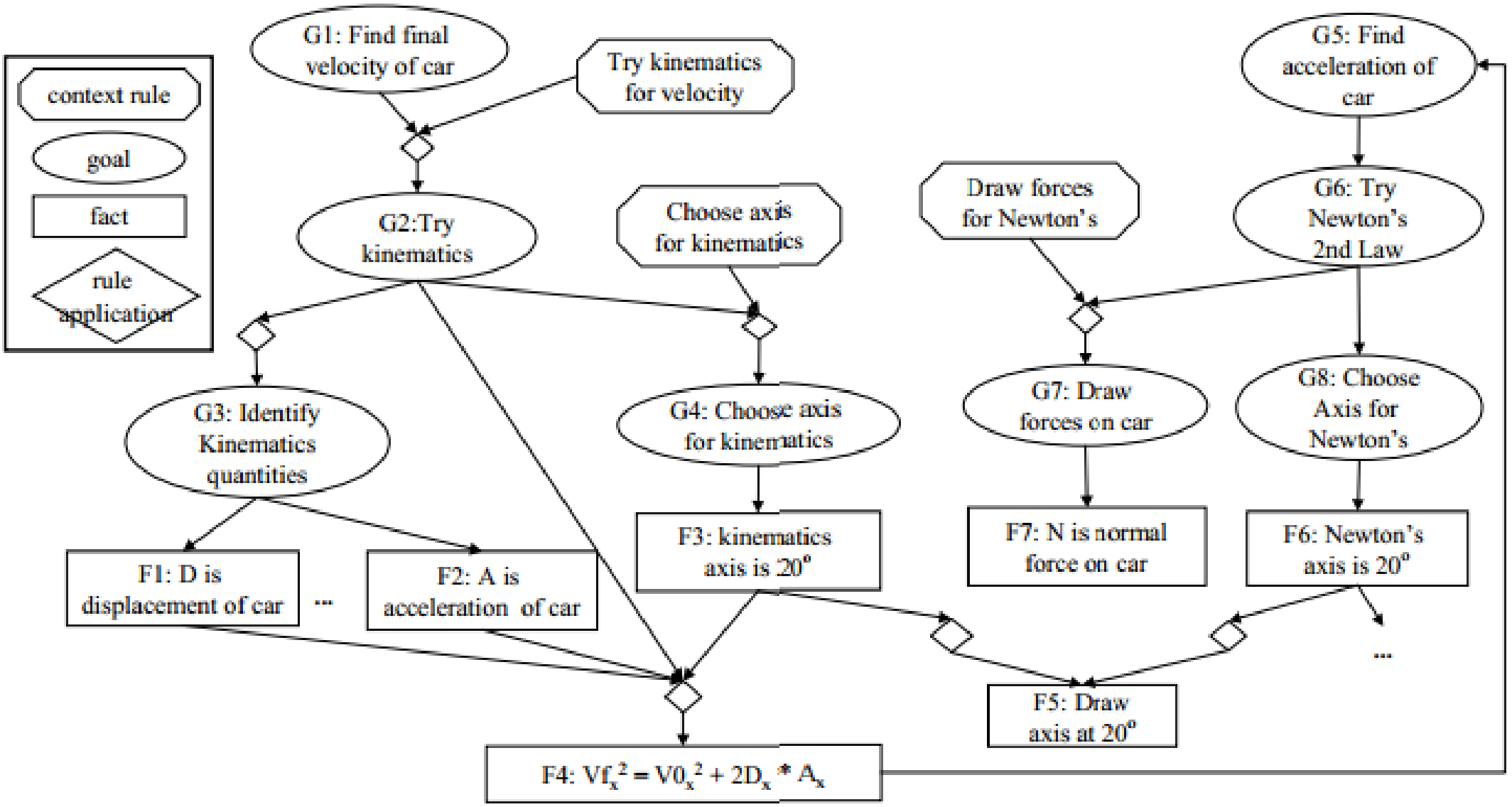
Use of Bayesian Models and Markov Models in Intelligent Tutoring Systems

Spyros Psychogios, Greta Kaufeld and Andy Russell

Bayesian Models

- Working definition:
- "The Bayesian network is a graphical modeling tool for specifying probability distributions" (Darwiche 2009).
- Why?
  - In order to handle uncertainty, for example concerning students' domain knowledge or problem solving strategy.
- Where?
  - For example, in Andes: Bayesian network is used to do long-term assessment of the student's domain knowledge, plan recognition (inferring the most likely strategy the student is using to solve a problem) and predictions of students' goals and actions during problem solving
- How?
  - Andes constructs its Bayesian networks from the output of a problem solver that generates all the acceptable solutions to a problem. The Bayesian network is then used to infer which part of the solution the student is working on and where she got stuck (Gertner et al. 1998)



A solution graph segment for a problem

Markov Models

- *Working Definition*: Markov Models are probabilistic Finite State Machines that consist of a set of states connected by transitions. As suggested by the Markov property, each state depends only on the previous one (Renals & Hain, 2010). Hidden Markov Models are named that way because their state sequence is not observable: only the output is. Parameters: Transition probabilities between states, observation probabilities, and initial probabilities for each state (Jeong et al. 2008).

Hidden Markov Models

- Why?
  - because they give the researcher the opportunity to go beyond simply "counting" the occurrence of a phenomenon and allow them to find possible coherence between phenomena (Jeong et al .2008)
- Where?
  - for example, in Betty's Brain: Hidden Markov Models are used to find patterns in students' problem solving strategies
- How?
  - Jeong et al. (2008) used log files recorded during sessions with Betty's Brain: From these log files, six main activities were extracted (for example, editing the concept map or quizzing Betty) and HMMs were trained. Some states of the resulting HMMs consisted of only one student activity, while others contained multiple activities: From these, patterns were derived (for example, *map probing*, where students first edited the concept map and then asked Betty questions to see if the new concepts were taught correctly).

Hidden Markov Models

- Where else?
  - in AutoTutor: D'Mello and Graesser (2010) used HMMs to test a theory of cognitive disequilibrium
  - Beal, Mitra and Cohen (2003) try to estimate learner engagement using HMMs
- How?
  - D'Mello and Graesser (2010) trained a HMM using Baum-Welch re-estimation until parameters converged; it was found that the states of the HMM corresponded to components of the theory (for example, a "equilibrium state" emitted "flow", while a "disequilibrium state" emitted confusion)
- Why?
  - because student engagement is much more feasible to monitor with HMMs than with eye-tracking devices, as used in Affective AutoTutor

References

- Renals, S. and Hain, T. (2010): Speech Recognition. In Clark, A., Fox, C., and Lappin, S. (Eds.): The Handbook of Computational Linguistics and Natural Language Processing. Wiley-Blackwell.
- Jeong, H., Gupta, A., Roscoe, R., Wagster, J., Biswas, G., and Schwartz, D. (2008): Using hidden markov models to characterize students behaviors in learning-by-teaching environments. In: Lecture Notes in Computer Science 5091, pp. 614-625. Springer.
- D'Mello, S., and Graesser, A. (2010): Modeling Cognitive-Affective Dynamics with Hidden Markov Models. In R. Catrambone and S. Ohlsson (Eds.): Proceedings of the 32nd Annual Cognitive Science Society (pp. 2721-2726). Austin, TX: Cognitive Science Society.
- Beal, C., Mitra, S., and Cohen, P.R. (2007): Modeling learning patterns of students with a Tutoring System using Hidden Markov Models. In: Proceedings of the 13th International Conference on Artificial Intelligence in Education, pp. 238-245.
- Gertner A. S., Conati C., VanLehn K. (1998), Procedural help in Andes: Generating hints using a Bayesian network student. In: Proceedings of the 15th National Conference on Artificial Intelligence
- Darwiche, A. (2009): Modeling and Reasoning with Beyesian Networks. Cambridge: Cambridge University Press.