

Od prvopočátku herních algoritmů byla kombinace hledání a ohodnocení budoucích stavů hry základním konceptem. Turingův šachový algoritmus z roku 1950 dokázal myslet dva kroky vpřed a Shannonova práce z roku 1950 obsahuje rozsáhlou část o způsobech ohodnocení stavů, které mají být použity při vyhledávání. Samuelův program pro hru dáma z roku 1959 již kombinuje vyhledávání a ohodnocení které se učí prostřednictvím self-play a bootstrappingu. TD-Gammon na tyto myšlenky navazuje a využívá neuronové sítě k osvojení těchto komplexních funkcí. Kombinace hledávání a hodnocení stavů byla přítomna v mnoha milnících, kde počítače překonaly své lidské protějšky — DeepBlue pro šachy a AlphaGo pro Go.

Až donedávna byla tato kombinace omezena na hry s úplnou informací. Ale mnoho zajímavých problémů neposkytuje agentovi dokonalou informaci o prostředí. Tato práce uvádí čtenáře do metod umožňující hledání v hrách s neúplnou informací.