

Erklärbare Entscheidungen in twin-nativer autonomer Mobilität: Entscheidungsspuren im S-PACT-Zyklus

Stephan Pareigis¹

¹ HAW Hamburg, Fakultät Informatik und Digitale Gesellschaft, Deutschland

Kurzfassung Eine Fahrentscheidung ist nachträglich nur dann überprüfbar, wenn die Größen, auf denen sie beruht, zum Entscheidungszeitpunkt festgehalten werden. Für die twin-native Architektur S-PACT wird festgelegt, welche Größen das sind: Die Entscheidungsspur ist ein pro Entscheidungszeitpunkt protokolliertes Tupel aus semantischem Zustand des Digitalen Zwillings, den im Prädiktiven Zwilling simulierten Handlungskandidaten mit ihren prädizierten Verläufen, der ausgeführten Handlung und den Gründen für die Verwerfung der übrigen Kandidaten. Gefordert werden Vollständigkeit, Lokalität und Nachprüfbarkeit der Spur. Die Trennung von faktischem und prädiktivem Zwilling schafft dafür die strukturellen Voraussetzungen, erzeugt die Spur jedoch nicht; vorgeschlagen wird deshalb eine explizite Repräsentation von Kandidaten, prädizierten Verläufen und Verwerfungsgründen. Dargestellt wird der Vorschlag an einem Minimalaufbau der Miniature Autonomy und gegen den implementierten Funktionsumfang abgegrenzt. Als Prüfkriterien werden Vollständigkeitsrate, Rekonstruierbarkeit und Prognosefehler benannt; ihre Erhebung steht aus.

Schlüsselwörter— erklärbare Autonomie, Digitaler Zwilling, Prädiktiver Zwilling, neuro-symbolische Architektur, Entscheidungsspur, Miniature Autonomy

I. MOTIVATION UND FORSCHUNGSFRAGE

Autonome Fahrfunktionen werden zunehmend datengetrieben realisiert. Damit wächst der Abstand zwischen der ausgeführten Handlung und einer Begründung, die ein Mensch nachvollziehen und ein Dritter überprüfen kann. Für Zulassung, Unfallanalyse und Akzeptanz ist nicht allein entscheidend, ob ein System sich korrekt verhält, sondern ob im Nachhinein angegeben werden kann, worauf eine einzelne Entscheidung beruhte.

Ob eine solche Angabe möglich ist, entscheidet der Architekturentwurf: Tragen kann eine Begründung

nur diejenigen Größen, die während der Entscheidung entstehen und festgehalten werden. Daraus ergibt sich die Frage: Welche Struktur muss eine twin-native Architektur aufweisen, damit zu einer einzelnen Fahrentscheidung eine vollständige und nachprüfbare Begründung angegeben werden kann?

Beantwortet wird die Frage konzeptionell. Vorgeschlagen wird mit der Entscheidungsspur ein operationaler Erklärungs-begriff; gezeigt wird, welche Bestandteile der Architektur S-PACT zu dessen Bestandteilen beitragen, und benannt werden Kriterien, an denen sich der Vorschlag künftig messen lassen muss. Architektur und Verarbeitungszyklus sind in [1] eingeführt und werden hier nicht als eigener Beitrag beansprucht; neu sind der Erklärungs-begriff, die daraus abgeleitete Repräsentationsschicht und die Prüfkriterien. Neue empirische Ergebnisse werden nicht berichtet.

II. FORSCHUNGSSTAND

Erklärungen für Fahrentscheidungen werden bislang überwiegend nachträglich erzeugt. Überblicksarbeiten unterscheiden Adressaten, Zeitpunkte und Formate solcher Erklärungen [4] und ordnen die Verfahren nach ihrem Ansatzpunkt im Verarbeitungsstapel [5]; für bildbasierte Fahrsysteme liegt eine systematische Übersicht ihrer Grenzen vor [6]. Ein großer Teil dieser Verfahren rekonstruiert aus einem bereits trainierten Modell nachträglich eine Begründung, etwa über Saliency-Karten oder sprachliche Beschreibungen; daneben werden auch von vornherein interpretierbare Entwürfe untersucht [6]. Für die rekonstruierenden Verfahren gilt, dass ihre Übereinstimmung mit dem tatsächlichen Zustandekommen der Entscheidung selbst wieder eine offene Prüffrage ist. Daraus ergibt sich eine Trennlinie: Eine Begründung kann nachträglich rekonstruiert oder während der Entscheidung mitgeschrieben werden.

Mitschreiben setzt voraus, dass die für eine Begründung nötigen Größen im System als benannte Objekte existieren. Diese Forderung ist außerhalb der Mobilität bereits formuliert worden: Unter dem

Begriff der decision provenance wird vorgeschlagen, Entscheidungsketten samt Eingangsgrößen und Folgewirkungen systematisch aufzuzeichnen, um automatisierte Systeme rechenschaftsfähig zu machen [7]. Offen bleibt dort, welche Größen ein geschlossener Regelkreis in jedem einzelnen Zyklus bereitstellen muss, damit eine solche Aufzeichnung vollständig ist. Die Digital-Twin-Forschung stellt geeignete Größen bereit, verfolgt sie jedoch unter anderen Fragestellungen: Der Sim-to-Real-Transfer und seine Fehlerquellen sind umfassend beschrieben [8], kontinuierlich synchronisierte Zwillinge dienen der Bewertung realer Regelstrategien am mitlaufenden Modell [9]. Die Zwischenrepräsentationen, an die eine Begründung anknüpfen könnte, liefern neuro-symbolische Ansätze zur Kopplung von subsymbolischer Wahrnehmung und symbolischer Handlungsplanung [10] sowie die hierarchische Verstärkungslernliteratur mit ihrer Zerlegung von Verhalten in wiederverwendbare Teilfertigkeiten [11]; der zugrunde liegende modulare Zyklus geht auf die klassische deliberative Zerlegung in Wahrnehmung, Planung und Handlung zurück [12].

Offen bleibt in beiden Strängen dieselbe Frage: Welche dieser Größen machen zusammengenommen eine Begründung aus, und wann ist sie vollständig? Solange das nicht festgelegt ist, bleibt der Digitale Zwilling ein Mittel zur Verbesserung von Übertragbarkeit und Sicherheit, aber kein Träger einer Begründungsstruktur. Genau diese Festlegung wird im Folgenden vorgenommen.

III. HINTERGRUND: S-PACT UND DER SUPPA-ZYKLUS

S-PACT (Synchronized Predictive Adaptive Collaborative Twin) ist eine neuro-symbolische Architektur, die neuronale Wahrnehmung, symbolische Repräsentation und zwei getrennte Zwillinge verbindet [1]. Der Digitale Zwilling (DT) hält den beobachtbaren Zustand des realen Systems in strukturierter, semantisch annotierter Form. Er simuliert nicht und prognostiziert nicht, sondern dient als faktische Verankerung mit nachvollziehbarer Datenherkunft. Der Prädiktive Zwilling (PT) ergänzt den DT um zeitliche und kausale Modelle. In ihm werden Zustandsübergänge ausgeführt, Handlungsalternativen durchgerechnet und Regelstrategien erlernt, bevor sie auf dem realen System zum Einsatz kommen. Beide Zwillinge werden fortlaufend abgeglichen.

Der Ablauf der Verarbeitung wird in [1] als SUPPA-Zyklus beschrieben, eine Erweiterung der klassischen deliberativen Zerlegung in Wahrnehmung, Planung und Handlung [12] um zwei Stufen. Die Stufen sind als funktionale Zuordnung zu verstehen und nicht als starre Verarbeitungsreihenfolge; die Kritik an einer solchen Festschreibung ist seit

den verhaltensbasierten Architekturen bekannt [13]. Sense überführt Rohdaten in strukturierte Beobachtungen. Understand bildet diese auf die semantische Repräsentation des DT ab, also auf Objekte, Relationen und Rollen. Predict rechnet im PT die Folgen möglicher Handlungen über einen endlichen Horizont durch. Plan wählt anhand dieser Prognosen eine Handlung aus. Act führt sie aus und speist die Abweichung zwischen Prognose und Messung in den nächsten Zyklus zurück.

IV. ENTSCHEIDUNGSSPUREN ALS ERKLÄRUNGSBEGRIFF

Als Erklärung einer Fahrentscheidung zum Zeitpunkt t wird im Folgenden nicht eine nachträgliche Zuschreibung verstanden, sondern ein während der Entscheidung entstehendes Protokoll. Die Entscheidungsspur $E(t)$ ist definiert als das Tupel

$$E(t) = (z(t), A(t), P(t), a^*(t), R(t))$$

mit dem semantischen Zustand $z(t)$ im DT, der Menge $A(t)$ der zum Zeitpunkt t zulässigen Handlungskandidaten, den im PT für diese Kandidaten prädizierten Verläufen $P(t)$, der ausgeführten Handlung $a^*(t)$ und den Verwerfungsgründen $R(t)$, also der Angabe, welches Kriterium die nicht gewählten Kandidaten verletzt haben.

An diesen Begriff werden drei Anforderungen gestellt. Vollständigkeit verlangt, dass alle fünf Bestandteile für jede protokollierte Entscheidung vorliegen. Lokalität verlangt, dass sich die Spur auf einen einzelnen Entscheidungszeitpunkt bezieht und nicht auf ein aggregiertes Verhalten über eine Fahrt. Nachprüfbarkeit verlangt, dass jeder Bestandteil unabhängig gegen aufgezeichnete Daten geprüft werden kann: $z(t)$ gegen die Messung, $P(t)$ gegen den tatsächlich eingetretenen Verlauf und $a^*(t)$ gegen die im Fahrzeug ausgeführten Stellgrößen.

Die zentrale These lautet: Die in S-PACT vorgenommene Trennung von faktischem DT und kausal-zeitlichem PT ist eine besonders geeignete, aber weder eine notwendige noch eine hinreichende Struktur für $E(t)$. Nicht notwendig ist sie, weil andere Architekturen dieselben Bestandteile explizit repräsentieren können, etwa modellprädiktive Regelung oder modellbasiertes Verstärkungslernen mit expliziten Rollouts; notwendig ist allein, dass die fünf Bestandteile überhaupt als benannte Objekte vorliegen. Geeignet ist sie, weil $z(t)$ und $P(t)$ in ihr verschiedene Herkunft und verschiedene Prüfinstanzen besitzen: Der DT ist gegen die Messung prüfbar, der PT gegen den tatsächlich eingetretenen Verlauf. Damit zerfällt die Nachprüfbarkeit der Spur in zwei unabhängige Prüfungen statt in eine gemeinsame. Nicht hinreichend ist die Trennung, weil der PT in [1] als Simulations- und Lernumgebung eingeführt ist: Er stellt die Mittel bereit,

Handlungsalternativen durchzurechnen, erzeugt jedoch im laufenden Betrieb nicht von selbst zu jeder ausgeführten Handlung eine protokollierte Kandidatenmenge samt Verwerfungsgründen.

Abzugrenzen ist der Vorschlag gegen rein end-to-end gelernte Regler ohne explizite Kandidaten- und Bewertungsrepräsentation. Sie liefern $a^*(t)$ und implizit $z(t)$, aber weder $A(t)$ noch $P(t)$ noch $R(t)$, sodass Begründungen dort nachträglich approximiert werden müssen. Wertbasierte Verfahren über einer diskreten Aktionsmenge stellen dagegen $A(t)$ und eine Rangfolge bereit; was ihnen fehlt, ist $R(t)$ im hier verwendeten Sinn, also die Angabe des verletzten Kriteriums statt eines bloßen Wertunterschieds.

Vorgeschlagen wird deshalb eine zusätzliche Schicht über der bestehenden Architektur: die explizite Repräsentation von $A(t)$, $P(t)$ und $R(t)$ im Entscheidungszyklus sowie deren Protokollierung.

Der kritische Bestandteil ist damit $R(t)$. Verwerfungsgründe lassen sich nur angeben, wenn die Plan-Stufe eine Auswahl über einer benannten Kandidatenmenge trifft und ihr Auswahlkriterium explizit vorliegt. Die vorgeschlagene Schicht ist damit nicht kostenlos, sondern schränkt die Gestaltung der Plan-Stufe ein: Sie verlangt entweder eine Kandidatengenerierung mit anschließender Bewertung im PT oder eine gelernte Strategie, deren Ausgabe über einer endlichen Kandidatenmenge normiert und deren Bewertungsfunktion mitgeführt wird. Ob sich diese Einschränkung ohne Einbußen an Regelgüte einlösen lässt, ist offen und stellt die zentrale technische Frage des Vorschlags dar.

V. MINIMALAUFBAU: MINIATURE AUTONOMY

Als Untersuchungsgegenstand dient ein bewusst reduzierter Aufbau. In einer Modellwelt im Maßstab 1:87 fährt ein Fahrzeug mit Bordkamera und Mikrocontroller; zwei Deckenkameras erfassen fortlaufend seine Pose und aktualisieren damit den DT [2]. Die Fahrbahn ist im DT als zweidimensionales Belegungsgitter mit Mittellinien und Rändern abgebildet, das Fahrzeug durch ein kinematisches Einspurmodell mit dem Lenkwinkel als einziger Stellgröße. Synthetische Kamerabilder werden mit den Parametern der realen Kamera gerendert, sodass beide Domänen denselben semantischen Beobachtungsraum besitzen. Die Querabweichung von der Mittellinie wird in beiden Domänen identisch berechnet und dient als Abgleichsignal; bei größerer Divergenz hält der PT die Vorwärtssimulation an und reinitialisiert seinen Zustand aus dem DT [1].

δ bezeichnet im Folgenden einen Lenkwinkelkandidaten, H den endlichen Prädiktionshorizont und CTE (cross-track error) die Querabweichung des Fahrzeugs von der Fahrspurmitte. Abb. 1 zeigt den

Zusammenhang an einem einzelnen Entscheidungszeitpunkt t_0 .

Entscheidungszeitpunkt t_0 : Annäherung an eine Kurve

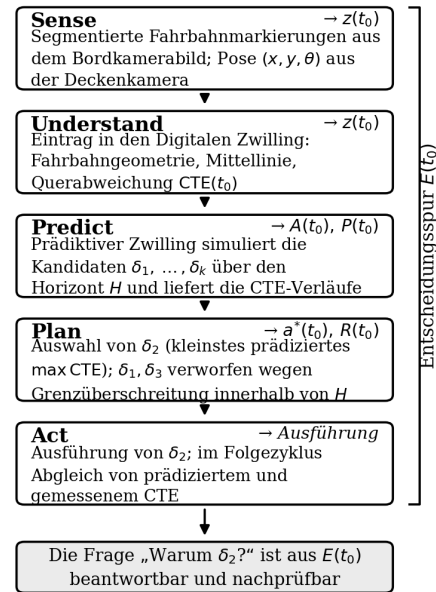


Abb. 1 Zerlegung einer einzelnen Fahrentscheidung in die Bestandteile der Entscheidungsspur, dargestellt am Fahrscenario der Miniature Autonomy (Abschnitt 5). Von oben nach unten sind die fünf Stufen des SUPPA-Zyklus aufgetragen; die kursive Angabe am rechten Rand jedes Kastens nennt den Bestandteil von $E(t_0)$, den die jeweilige Stufe erzeugt. Der Darstellung ist zu entnehmen, dass die Spur kein zusätzliches Erhebungsverfahren erfordert: Sense und Understand liefern gemeinsam $z(t_0)$, Predict die Kandidatenmenge $A(t_0)$ samt prädizierten Verläufen $P(t_0)$, Plan die ausgeführte Handlung $a^*(t_0)$ und die Verwerfungsgründe $R(t_0)$. Alle fünf Bestandteile entstehen an genau einer Stelle des ohnehin durchlaufenen Zyklus, sodass die Frage nach der Begründung einer Stellgröße ohne nachträgliche Rekonstruktion beantwortbar ist. Eigene Darstellung.

Dieser Aufbau folgt dem Prinzip der minimalen Modellierung: Es wird nicht versucht, die Realität möglichst vollständig abzubilden, sondern nur diejenigen Eigenschaften, die für die untersuchte Frage entscheidend sind. Für den Erklärungs-begriff ist das ein Vorteil, denn der Zustand $z(t)$ besteht aus wenigen, einzeln prüfbaren Größen: Pose, Fahrbahngeometrie und Querabweichung. Damit ist die Nachprüfbarkeit des Zustandsanteils $z(t)$ nicht nur formulierbar, sondern praktisch einlösbar; für $P(t)$ und $R(t)$ gilt das noch nicht.

Der in Abb. 1 dargestellte Ablauf ist entsprechend zu lesen: Er beschreibt, wie sich eine Entscheidungsspur mit den vorhandenen Komponenten protokollieren ließe, nicht ein bereits durchgeführtes Experiment. Tabelle 1 grenzt ab, welche Bestandteile heute implementiert sind und welche den vorgeschlagenen Ausbau darstellen.

Bestandteil	Stand
Synchronisierter DT: Pose und Fahr- bahngeometrie [2]	implementiert
Identische Berechnung der Querab- weichung in beiden Domänen [1]	implementiert
Reinitialisierung des PT aus dem DT bei Divergenz [1]	implementiert
Erlernen und Rückspielen von Regel- strategien, Real2Sim2Real [1]	implementiert
Prädiktive Aufgabenvalidierung am Manipulator [3]	implementiert, an- derer Demonstra- tor
Explizite Kandidatenmenge $A(t)$ und Bewertung $P(t)$ je Zyklus	vorgeschlagen
Protokollierung vollständiger Entschei- dungsspuren $E(t)$	vorgeschlagen
Sprachliche Aufbereitung der Spur für Adressaten [4]	vorgeschlagen

TABELLE 1 Abgrenzung des implementierten Standes von der vorgeschlagenen Erweiterung. Die linke Spalte nennt den Bestandteil, die rechte seinen Realisierungsstand; Quellenangaben verweisen auf die Arbeit, in der der jeweilige Bestandteil beschrieben ist. Der Tabelle ist zu entnehmen, dass die Zwillingssynchronisation und der Lernkreislauf bereits realisiert sind, während genau die drei Bestandteile fehlen, die die Entscheidungsspur ausmachen: die explizite Kandidatenmenge, deren Bewertung und die Protokollierung. Die vorgeschlagene Erweiterung betrifft ausschließlich die untere Gruppe.

VI. LIMITATIONEN UND GEPLANTE EVALUATION

Nicht belegt wird, dass Entscheidungsspuren Menschen tatsächlich beim Verstehen helfen; dies wäre in einer Nutzerstudie zu prüfen und ist eine von der Struktur unabhängige Frage. Ebenso wenig wird gezeigt, dass die vorgeschlagene Struktur post-hoc erzeugten Erklärungen überlegen ist, da ein Vergleich mit entsprechenden Verfahren aussteht. Eine protokollierte Spur begründet eine Entscheidung, sie rechtfertigt sie nicht: Ist das Prognosemodell im PT falsch, so ist auch die Begründung nachvollziehbar falsch. Der betrachtete Aufbau ist zudem einspurig, einzelfahrzeugbezogen und frei von anderen Verkehrsteilnehmern, und die in Abschnitt 4 benannte Einschränkung der Plan-Stufe ist ein Eingriff in den Regelentwurf, dessen Auswirkung auf die Regelgüte hier nicht bestimmt wird.

Für die Erweiterung sind drei Größen vorgesehen, die sich ohne Nutzerbeteiligung erheben lassen. Die Vollständigkeitsrate gibt den Anteil der Zyklen an, für die alle fünf Bestandteile von $E(t)$ vorliegen. Die Rekonstruierbarkeit prüft, ob eine erneute Simulation aus der aufgezeichneten Spur dieselbe Handlung innerhalb einer Toleranz reproduziert. Der Prognosefehler misst die Abweichung zwischen prädizierter und gemessener Querabweichung über den Horizont und bewertet damit, wie belastbar der Bestandteil $P(t)$ ist. Erst diese drei Größen erlauben es, die eingangs gestellte Frage empirisch zu beantworten.

VII. FAZIT

Vorgeschlagen wurde, Erklärbarkeit in twin-nativer autonomer Mobilität nicht als Eigenschaft eines Modells, sondern als Protokolleigenschaft eines Entscheidungszyklus zu fassen. Die Entscheidungsspur benennt fünf Bestandteile, die eine Architektur erzeugen muss. Die Trennung zwischen faktischem und prädiktivem Zwilling in S-PACT schafft dafür die Voraussetzungen, ersetzt aber nicht die vorgeschlagene explizite Repräsentation von Handlungskandidaten, prädizierten Verläufen und Verwerfungsgründen. Sein Wert entscheidet sich an der Frage, ob sich vollständige, rekonstruierbare Spuren im laufenden Betrieb des beschriebenen Aufbaus erheben lassen. Die dafür nötigen Kriterien sind benannt, ihre Erhebung ist der nächste Schritt.

LITERATUR

- [1] S. Pareigis and H. Voss, "A neuro-symbolic architecture for twin-native autonomy," in *Advances in Real-Time and Autonomous Systems (AUTSYS 2025)*, H. Unger and M. Schaible, Eds., *Lecture Notes in Networks and Systems*, vol. 1979. Cham, Switzerland: Springer, 2026, pp. 53–63.
- [2] S. Pareigis, D. L. Riege, and T. Tiedemann, "Miniature autonomous vehicle environment for sim-to-real transfer in reinforcement learning," in *Proc. 21st Int. Conf. Informatics in Control, Automation and Robotics (ICINCO)*, vol. 1, 2024, pp. 309–317.
- [3] L. Jedelhauser, "PACT: A framework for predictive task validation and adaptive error recovery in robotic systems," M.S. thesis, HAW Hamburg, Hamburg, Germany, 2025.
- [4] D. Omeiza, H. Webb, M. Jirotko, and L. Kunze, "Explanations in autonomous driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10142–10162, Aug. 2022.
- [5] S. Atakishiyev, M. Salameh, H. Yao, and R. Goebel, "Explainable artificial intelligence for autonomous driving: A comprehensive overview and field guide for future research directions," *IEEE Access*, vol. 12, pp. 101603–101625, 2024.
- [6] É. Zablocki, H. Ben-Younes, P. Pérez, and M. Cord, "Explainability of deep vision-based autonomous driving systems: Review and challenges," *Int. J. Comput. Vis.*, vol. 130, no. 10, pp. 2425–2452, 2022.
- [7] J. Singh, J. Cobbe, and C. Norval, "Decision provenance: Harnessing data flow for accountable systems," *IEEE Access*, vol. 7, pp. 6562–6574, 2019.
- [8] W. Zhao, J. P. Queralta, and T. Westerlund, "Sim-to-real transfer in deep reinforcement learning for robotics: A survey," in *Proc. IEEE Symp. Series*

Computational Intelligence (SSCI), 2020, pp. 737–744.

- [9] J. Abou-Chakra et al., “Real-is-sim: Bridging the sim-to-real gap with a dynamic digital twin,” arXiv:2504.03597v2, 2025.
- [10] M. Eppe, P. D. H. Nguyen, and S. Wermter, “From semantics to execution: Integrating action planning with reinforcement learning for robotic causal problem-solving,” *Front. Robot. AI*, vol. 6, art. 123, 2019.
- [11] R. S. Sutton, D. Precup, and S. Singh, “Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning,” *Artif. Intell.*, vol. 112, no. 1–2, pp. 181–211, 1999.
- [12] R. E. Fikes and N. J. Nilsson, “STRIPS: A new approach to the application of theorem proving to problem solving,” *Artif. Intell.*, vol. 2, no. 3–4, pp. 189–208, 1971.
- [13] R. A. Brooks, “A robust layered control system for a mobile robot,” *IEEE J. Robot. Autom.*, vol. 2, no. 1, pp. 14–23, 1986.